

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 15:59:37 ; Search time 36 Seconds

(without alignments)
495,988 Million cell updates/sec

Title: US-10-048-196-2

Perfect score: 676
Sequence: 1 NMKILYVATLMTAFITLASC.....SLKPCMLETVNMFVPTTR 134

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	676	100.0	134	22	AAB60646
2	118	17.5	173	21	AA174455
3	116	17.2	173	21	AA174454
4	110.5	16.3	172	21	AA174453
5	82	12.1	179	23	ABBS5176
6	81	12.0	202	22	AAU55473
7	79.5	11.8	4274	22	ABG00972
8	79.5	11.8	4386	22	ABG07375
9	79.5	11.8	4397	22	ABG21944
10	79	11.7	194	22	AAU65361

11	78.5	11.6	200	22	AAG92091	C glutamicum prote
12	78	11.5	579	22	AAG93285	Human protein HPI0
13	78	11.5	580	23	ABBS97409	Novel human protei
14	77.5	11.5	764	23	ABBS05596	Human testis devel
15	77.5	11.5	789	22	ABBS9542	Human protein sequ
16	76.5	11.3	386	22	ABBS29832	Peptide #2483 enco
17	76.5	11.3	386	22	ABBS5008	Peptide #2514 enco
18	76.5	11.3	386	22	ABBS20422	Protein #2421 enco
19	76.5	11.3	386	22	AAAS5821	Human brain expres
20	76.5	11.3	386	22	AAAB68195	Human bone marrow
21	76.5	11.3	386	22	AAAM16015	Human bone marrow
22	76.5	11.3	386	22	AAAM28516	Peptide #2553 enco
23	76.5	11.3	386	22	AAAM03749	Peptide #2431 enco
24	76	11.2	310	21	AAAG1397	Arabidopsis thalia
25	76	11.2	357	21	AAAG1396	Arabidopsis thalia
26	76	11.2	515	21	AAAG1395	Arabidopsis thalia
27	75	11.1	1203	21	AAAY83275	Candida albicans C
28	75	11.1	1203	21	AAAY84814	Amino acid sequenc
29	74	10.9	298	14	AAAR9286	Respiratory syncyt
30	74	10.9	298	14	AAAR9286	Membrane bound G p
31	73.5	10.9	878	20	AAEL14719	Human carboxylate
32	73	10.8	89	22	AAU57395	Propionibacterium
33	73	10.8	197	22	AAU57478	Propionibacterium
34	72.5	10.7	160	21	AAAY75512	Neisseria meningit
35	72.5	10.7	436	21	AAAG18612	Arabidopsis thalia
36	72.5	10.7	436	21	AAAG18612	Arabidopsis thalia
37	72.5	10.7	544	21	AAAG18611	Arabidopsis thalia
38	72.5	10.7	544	21	AAAG18611	Arabidopsis thalia
39	72.5	10.7	584	21	AAAG18610	Arabidopsis thalia
40	72.5	10.7	584	21	AAAG18610	Arabidopsis thalia
41	72	10.7	1935	22	ABBS9858	Drosophila melanog
42	72	10.7	2368	22	AAU34139	Staphylococcus aur
43	72	10.7	2368	22	AAU34139	Staphylococcus aur
44	71.5	10.6	737	21	AAAB59023	Breast and ovarian
45	71.5	10.6	974	22	ABBS62642	Drosophila melanog

ALIGNMENTS

RESULT 1
AAB60646
ID AAB60646 standard; Protein: 134 AA.
XX
AC AAB60646;
XX
DT 04-MAY-2001 (first entry)
XX
DE Moraxella catarrhalis strain ATCC43617 BASB125 protein.
XX
KW BASB125 protein; strain ATCC43617; antigen; antibody; vaccine;
KW genetic immunisation; infection; upper respiratory tract; otitis media;
KW hearing loss; deafness; pneumonia; sinusitis; nosocomial infection;
KW invasive disease; antibacterial; auditory.
XX
OS Moraxella catarrhalis.
XX
PN W0200109331-A2.
XX
PD 08-FEB-2001.
XX
PF 27-JUL-2000; 2000MO-EF07291.
XX
PR 30-JUL-1999; 99GB-0018041.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Thonard J;
XX
DR WPI: 2001-168707/17.
XX
PT N-PSDB; AAF59800.
XX
New BASB125 polypeptide isolated from Moraxella catarrhalis for

PT treating, preventing and diagnosing diseases associated with M.
 PT catarrhalis infection in mammals, e.g. otitis media in humans
 XX
 PS Claim 4; Page 64; 73pp; English.
 CC
 CC The invention relates to the Moraxella catarrhalis strain ATCC43617
 CC BASB125 protein (AA80646) and to DNA encoding it (AA59800). The
 CC invention also relates to immunogenic fragments of the BASB125 protein,
 CC expression vectors and host cells comprising BASB125 nucleic acids, the
 CC recombinant production of BASB125, vaccine compositions comprising the
 CC BASB125 protein or nucleic acid, an antibody against BASB125, therapeutic
 CC compositions comprising the anti-BASB125 antibody, and a method of
 CC identifying a Moraxella catarrhalis infection via the detection of
 CC BASB125 proteins or antibodies. The vaccine compositions of the invention
 CC are useful as prophylactic or therapeutic agents against Moraxella
 CC catarrhalis infections in mammals, particularly humans. Moraxella
 CC catarrhalis is a Gram negative bacterium frequently isolated from the
 CC human upper respiratory tract, which is responsible for several
 CC pathological conditions. It is responsible for about 15% of otitis media
 CC cases in children (which can lead to temporary or permanent hearing
 CC loss). It also causes pneumonia in elderly people, and sinusitis,
 CC nosocomial infections and, less frequently, invasive diseases. BASB125
 CC proteins or nucleotides may additionally be used in screening for novel
 CC antibacterial compounds, and in the diagnosis and staging of infections.
 CC The present sequence represents the Moraxella catarrhalis strain
 CC ATCC43617 BASB125 protein.
 CC
 XX
 SQ Sequence 134 AA;
 Query Match 100.0%; Score 676; DB 22; Length 134;
 Best Local Similarity 100.0%; Pred. No. 5.5e-73;
 Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKKILVYATATLMTAFTLASCASPESNPKNSSANLTTSLIKHAKQTCOTLTGHQYWKI 60
 DB 1 MKKILVYATATLMTAFTLASCASPESNPKNSSANLTTSLIKHAKQTCOTLTGHQYWKI 60
 QY 61 AAKKLSSESRAKISSETACGCVADKAPPAVSLTETLTAINPAREVAOKIVRHSLSKP 120
 DB 61 AAKKLSSESRAKISSETACGCVADKAPPAVSLTETLTAINPAREVAOKIVRHSLSKP 120
 QY 121 LETVNAFIYPTTTR 134
 DB 121 LETVNAFIYPTTTR 134
 AC AAY74455;
 XX
 DT 21-MAR-2000 (first entry)
 DE Neisseria meningitidis ORF 108 protein sequence SEQ ID NO:386.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
 KW antibacterial; gene therapy.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO9957280-A2.
 PD 11-NOV-1999.
 PE 30-APR-1999; 99WO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.

PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX
 PA (CHIR) CHIRON CORP.
 PI (GENO) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;
 PI Petersen J, Pizzo M, Rappold R, Ratti G, Scialto E, Scarselli M;
 PI Tettelin H, Venter JC;
 DR WPI: 2000-062150/05.
 DR N-PsDB: AA253217.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics
 PS Claim 2; Page 320; 1453pp; English.
 CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC presence of Neisseria bacteria (e.g. meningitis and septicaemia), to detect the
 CC be used to screen for agonists or antagonists, which may also
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 CC
 XX
 SQ Sequence 173 AA;
 Query Match 17.5%; Score 118; DB 21; Length 173;
 Best Local Similarity 23.1%; Pred. No. 7.4e-06;
 Matches 28; Conservative 29; Mismatches 58; Indels 6; Gaps 3;
 QY 1 MKKILVYATATLMTAFTLASCASPES--NPKNSANLTTSLIKHAKQTCOTLTGHQY 58
 DB 51 MKKTLST---LPVAIILGGCGAGGNTFGSLDGTGSGSIVKMAVBSQCAELNKRSEW 107
 QY 59 KIAAMKLSSESRAKISSETACGCVADKAPPAVSLTETLTAINPAREVAOKIVRHSLSKP 118
 DB 108 RLTAIAWSAKQAEWEKIKACVAGQEPNDITGNDV-MQMLDPSYRNQALALAKTVSA 166
 QY 119 C 119
 DB 167 C 167
 AC AAY74454;
 XX
 DT 21-MAR-2000 (first entry)
 DE Neisseria meningitidis ORF 108 protein sequence SEQ ID NO:384.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
 KW antibacterial; gene therapy.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO9957280-A2.
 PD 11-NOV-1999.
 PE 30-APR-1999; 99WO-US09346.

XX 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR WPI: 2000-062150/05.
 DR N-PSDB; AAZ53216.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics -
 XX
 PS Claim 2; Page 319; 1453pp; English.
 XX
 CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
 CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC *Neisseria meningitidis* (e.g. meningitis and septicaemia), to detect the
 CC presence of *Neisseria meningitidis*, or to raise antibodies. They may also
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 CC
 XX
 SQ Sequence 173 AA;
 Query Match 17.2%; Score 116; DB 21; Length 173;
 Best Local Similarity 24.0%; Pred. No. 1.3e-05;
 Matches 29; Conservative 27; Mismatches 59; Indels 6; Gaps 3;
 XX
 QY 1 MKILYVATMTAFTATASCSTPES--NPKSSANLTTSLIKHAVKQCTQTLGHQY 58
 Db 51 MKKTLST---LPVAILLGCGCAAGGNTFGSLDGGTGKGGSTIVKMAVGSQCRALDRKSEW 107
 QY 59 KIAAMKLSSESKAKISSETACGCVADKAPAEVSLTETTTAANPNARTEVAOKIVRHSIKP 118
 Db 108 RLTLAAMSAEKQAEWENKICGCAVQAEAPERMTGNDV-MQMLAPSTRNALALATKATYSA 166
 QY 119 C 119
 Db 167 C 167
 XX
 RESULT 4
 ID AAY74453 standard; Protein; 172 AA.
 AC AAY74453;
 XX
 DT 21-MAR-2000 (first entry)
 XX
 DE *Neisseria gonorrhoeae* ORF 108 protein sequence SEQ ID NO:382.
 XX
 KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
 KM antigenic; diagnostic; immunogenic; infection; meningitis; septicaemia;
 XX antibacterial; gene therapy.
 OS *Neisseria gonorrhoeae*.

XX W09957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR WPI: 2000-062150/05.
 DR N-PSDB; AAZ53215.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics -
 XX
 PS Claim 2; Page 319; 1453pp; English.
 XX
 CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
 CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC *Neisseria meningitidis* (e.g. meningitis and septicaemia), to detect the
 CC presence of *Neisseria meningitidis*, or to raise antibodies. They may also
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 CC
 XX
 SQ Sequence 172 AA;
 Query Match 16.3%; Score 110.5; DB 21; Length 172;
 Best Local Similarity 21.7%; Pred. No. 5.9e-05;
 Matches 26; Conservative 29; Mismatches 60; Indels 5; Gaps 3;
 XX
 QY 1 MKKILYVATMTAFTATASCSTPES--NPKSSANLTTSLIKHAVKQCTQTLGHQY 59
 Db 51 MKKTLST---LPVAILLGCGCAAGGNTFGSLDGGTGKGGSTIVKMAVGSQCRALDRKSEW 107
 QY 60 IAAMKLSSESKAKISSETACGCVADKAPAEVSLTETTTAANPNARTEVAOKIVRHSIKP 119
 Db 108 RLTLAAMSAEKQAEWENKICGCAVQAEAPERMTGNDV-MQMLAPSTRNALALATKATYSA 166
 XX
 RESULT 5
 ID ABB55176 standard; Protein; 179 AA.
 AC ABB55176;
 XX
 DT 16-MAY-2002 (first entry)
 XX
 DE *Lactococcus lactis* protein ysfF.
 XX
 KW *Lactococcus lactis*; biodegradation; lactic bacterium; yogurt; cheese.
 XX
 OS *Lactococcus lactis* IL1403.

```

XX  FR2807446-A1.
XX
XX  12-OCT-2001.
XX
XX  11-APR-2000; 2000FR-0004630.
XX
XX  11-APR-2000; 2000FR-0004630.
XX
XX  (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX
XX  Bolotine A, Sorokine A, Renault P, Ehrlich SD;
XX  WPI: 2002-043418/06.
XX
XX  New nucleotide sequence useful in the identification of Lactococcus
XX  lactis and related species -
XX
XX  Claim 6; SEQ ID NO 1878; 2504bp; French.
XX
XX  The present invention is related to a Lactococcus lactis nucleotide
XX  sequence (ABA90521) and related proteins (ABB53300-ABB5621). The
XX  nucleic acid sequence is useful in the detection and/or amplification of
XX  nucleic acid sequence, particularly to identify Lactococcus lactis or
XX  related species. The proteins of the invention are useful for the
XX  biosynthesis or biodegradation of a composition of interest. The
XX  invention helps research in lactic bacteria, particularly useful in the
XX  production of yogurt and cheese.
XX  Note: The sequence data for this patent is based on equivalent patent
XX  WO2001/7734 (published 18-Oct-2001) which is available in electronic
XX  format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX  Sequence 179 AA;
XX
XX  Query Match 12.1%; Score 82; DB 23; Length 179;
XX  Best Local Similarity 23.1%; Pred. No. 0.16;
XX  Matches 33; Conservative 21; Mismatches 55; Indels 34; Gaps 4;
XX
XX  1 MKLITVYATATMTAFTTASCASPESNPKNSANLTTSLIKHAKQTCOTOLGHOYKI 60
XX  1 MKRLITTTTALALLSLGACSKSDASHKSSSTSSFPSPSTSS----- 45
XX
XX  61 AAMKLSSESRAKISSETACGCVADKAPAVSLTETLTAIINP-----NAPTEVAOKI 111
XX  46 SSSKVVSSSSKVV---EPTGNIDSSFOKAVADATQPTVATMKETKYSDITATAEAPQTI 102
XX
XX  112 VRHSLKRCQMETVNAFVPTTR 134
XX  103 V-----YTYTYRNQITPTATK 118
XX
XX  Db
XX
XX  RESULT 6
XX  AAU55473
XX  ID AAU55473 standard; Protein; 202 AA.
XX
XX  AAU55473;
XX
XX  27-FEB-2002 (first entry)
XX
XX  Propionibacterium acnes immunogenic protein #16369.
XX
XX  SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
XX  uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX  inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX  dermatological; osteopathic; neuroprotectant.
XX
XX  Propionibacterium acnes.
XX
XX  WO200181581-A2.
XX
XX  01-NOV-2001.
XX
XX  20-APR-2001; 2001WO-US12865.
XX

```

```

XX  21-APR-2000; 2000US-199047P.
XX
XX  02-JUN-2000; 2000US-20841P.
XX
XX  07-JUL-2000; 2000US-216747P.
XX
XX  (CORI-) CORIXA CORP.
XX
XX  Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX  L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX  WPI: 2001-616774/71.
XX
XX  N-PSDB; AAS59569.
XX
XX  Propionibacterium acnes polypeptides and nucleic acids useful for
XX  vaccinating against and diagnosing infections, especially useful for
XX  treating acne vulgaris -
XX
XX  Example 1; SEQ ID NO 16668; 1069pp; English.
XX
XX  Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX  polypeptides. The proteins and their associated DNA sequences are used in
XX  the treatment, prevention and diagnosis of medical conditions caused by
XX  P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX  pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
XX  P. acnes is also involved in infections of bone, joints and the central
XX  nervous system, however it is particularly involved in the inflammatory
XX  lesions associated with acne vulgaris. A method for detecting the
XX  presence or absence of P. acnes in a patient comprises contacting a
XX  sample with a binding agent that binds to the proteins of the invention
XX  and determining the amount of bound protein in the sample. The
XX  polypeptides may be used as antigens in the production of antibodies
XX  specific for P. acnes proteins. These antibodies can be used to
XX  downregulate expression and activity of P. acnes polypeptides and
XX  therefore treat P. acnes infections. The antibodies may also be used as
XX  diagnostic agents for determining P. acnes presence, for example, by
XX  enzyme linked immunosorbent assay (ELISA).
XX  Note: The sequence data for this patent did not form part of the printed
XX  specification, but was obtained in electronic format directly from WIPO
XX  at ftp.wipo.int/pub/published_pct_sequences.
XX
XX  Sequence 202 AA;
XX
XX  Query Match 12.0%; Score 81; DB 22; Length 202;
XX  Best Local Similarity 25.2%; Pred. No. 0.26;
XX  Matches 31; Conservative 27; Mismatches 41; Indels 24; Gaps 6;
XX
XX  12 MTAFTLASCASTPESNPK-----NSANLTTSLIKHAKQTCO--TOLGHOYKI 60
XX  25 LTIYRLKSSRESQRRASPIFWESSNSQMTGDIGHASVTGSCQLMSSVQVGEKMR 84
XX
XX  61 AAMKLSSESRAKISSETACGCVADKAPAVSLTETLTAIINPAPRT---EVAOKIV 112
XX  85 APMRI---AETKQOHVASGTRNLRLNPAKAPPSIAIS--LTSTIGPIARASMAVGEKLV 139
XX
XX  113 RHS 115
XX  140 RET 142
XX
XX  Db
XX
XX  RESULT 7
XX  ABG00972
XX  ID ABG00972 standard; Protein; 4274 AA.
XX
XX  ABG00972;
XX
XX  13-FEB-2002 (first entry)
XX
XX  Novel human diagnostic protein #963.
XX
XX  Human: chromosome mapping; gene mapping; gene therapy; forensic;
XX  food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX  Homo sapiens.
XX

```

XX WO200175067-A2.
PN 11-OCT-2001.
XX
XX
XX 30-MAR-2001; 2001WO-US08631.
PF
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
DR MPI; 2001-639362/73.
XX N-PSDB; AAS65159.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 31331; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4274 AA;
Query Match 11.8%; Score 79.5; DB 22; Length 4274;
Best Local Similarity 26.7%; Pred. No. 30;
Matches 28; Conservative 18; Mismatches 52; Indels 7; Gaps 4;
QY 14 AFLTASCASTPESNP-KNSSANLTSLIKHAVQOTQOTLTGHQYKIAMKLSSESKAK 72
DB 3900 ALTTSSCVDVKSRIPKNTPRD--NII--AVRKACATQOG-QPEKGAKOLPSKLPVK 3953
QY 73 ISETACGCVADKAPKPEAVSLTETLTAIINPAPTEVAOKIYRHSUK 117
DB 3954 VNSTCVTTTTTATTTTTTTTTTTTCTVAVKRSQLEKVKHSIE 3998
RESULT 8
ABG07375
ID ABG07375 standard; Protein: 4386 AA.
XX
AC ABG07375;
XX
XX 13-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #7366.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensics;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.
OS
XX
XX WO200175067-A2.
PN
XX
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001; 2001WO-US08631.
PE
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
DR MPI; 2001-639362/73.
XX N-PSDB; AAS71562.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 37734; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4386 AA;
Query Match 11.8%; Score 79.5; DB 22; Length 4386;
Best Local Similarity 26.7%; Pred. No. 31;
Matches 28; Conservative 18; Mismatches 52; Indels 7; Gaps 4;
QY 14 AFLTASCASTPESNP-KNSSANLTSLIKHAVQOTQOTLTGHQYKIAMKLSSESKAK 72
DB 3909 ALTTSSCVDVKSRIPKNTPRD--NII--AVRKACATQOG-QPEKGAKOLPSKLPVK 3962
QY 73 ISETACGCVADKAPKPEAVSLTETLTAIINPAPTEVAOKIYRHSUK 117
DB 3963 VNSTCVTTTTTATTTTTTTTTTCTVAVKRSQLEKVKHSIE 4007
RESULT 9
ABG21944
ID ABG21944 standard; Protein: 4397 AA.
XX
AC ABG21944;
XX
XX 18-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #21935.
XX

Human; chromosome mapping; gene mapping; gene therapy; forensic;
food supplement; medical imaging; diagnostic; genetic disorder.
Homo sapiens.
MO200175067-A2.
11-OCT-2001.
30-MAR-2001; 2001WO-US08631.
31-MAR-2000; 2000US-0540217.
23-AUG-2000; 2000US-0649167.
(HYSE-) HYSEQ INC.
Dmanac RT, Liu C, Tang YT;
WPI: 2001-639362/73.
N-PSDB: AAS86131.
New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity.
Claim 20; SEQ ID NO 52303; 103pp; English.
The invention relates to isolated polynucleotide (I) and
polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAG00010-ABG30377 represent novel human
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 4397 AA;
Query Match 11.8%; Score 79.5; DB 22; Length 4397;
Best Local Similarity 26.7%; Pred. No. 31;
Matches 28; Conservative 18; Mismatches 52; Indels 7; Gaps 4;
OY 14 AFTLASCASTPESNP-KNSSANLTSLIKHAVKOTCOTOLGHOYWKIAMLTSSSKAK 72
DB 3909 ALTTSSCVDKSRIPVKMTHRD---NIT--AVRKACATOKG-QPEKGAIAKOLPSKLPVK 3962
OY 73 ISETACGVADKAPAVSLTELTTAAMPNARTEVAQKIVRHSK 117
DB 3963 VRSTCVTTTTTATTTTSTTTTSCIVKVKRSQLEKCKHSIE 4007
RESULT 10
AAU65361
ID AAU65361 standard; Protein; 194 AA.
XX
AC AAU65361;
XX
DT 27-FEB-2002 (first entry)
XX

Propionibacterium acnes immunogenic protein #26257.
SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KM dermatological; osteopathic; neuroprotectant.
Propionibacterium acnes.
WO200181581-A2.
01-NOV-2001.
20-APR-2001; 2001WO-US12865.
21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
(CORI-) CORIXA CORP.
Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
WPI: 2001-616774/71.
N-PSDB: AAS59666.
Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
Example 1; SEQ ID NO 26556; 1069pp; English.
Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 194 AA;
Query Match 11.7%; Score 79; DB 22; Length 194;
Best Local Similarity 26.4%; Pred. No. 0.42;
Matches 39; Conservative 18; Mismatches 61; Indels 30; Gaps 6;
OY 9 AFLMTAFTLASCASTPESNPKNSSANL-----TSLIKHAVKOT---COTOLGHOYWKI 60
DB 25 STTSAACAPASANSMPPLPASTTRIDLSTSSILVESTSSPGCDLYESAQ----- 80
OY 61 AAMKLSSEKAKISTAGC-----CVADKAPAVSLTELTTAAMPNARTEVAQ 109
DB 81 CALDSTSPSVASVSAAPGSPVAIVLPASCVPKRTSTVSLDMNECGGYPEASTPPCH 140
OY 110 KIVRHSKLP-----CMLETVNAFIVPTT 133
DB 141 ---RAALRPRLRSLWOLPTVTAFLSPTTS 165

```
RESULT 11
AAG92091
ID AAG92091 standard; Protein: 200 AA.
XX
XX
AC AAG92091;
XX
XX
DT 26-SEP-2001 (first entry)
XX
XX
DE C glutamicum protein fragment SEQ ID NO: 5845.
XX
XX
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis.
XX
OS Corynebacterium glutamicum.
XX
PN EPI108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-0127688.
XX
PR 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
DR WPI: 2001-376931/40.
DR N-PSDB: AAH67310.
XX
XX
PT Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX
PS Claim 17; SEQ ID NO: 5845; 246pp + Sequence Listing; English.
XX
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 200 AA;
Query Match 11.6%; Score 78.5; DB 22; Length 200;
Best Local Similarity 23.5%; Pred. No. 0.5;
Matches 27; Conservative 28; Mismatches 47; Indels 13; Gaps 4;
OY 1 MKKILYATATMTAFTLASCASPESNPKNSSANLTSLIKHAVKOTCOTQLTGHOYWKI 60
:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3 LFKATAVTFVAALALISACSSSDSSSTSSSTSSASDAATQYPTAEELN----- 56
OY 61 AAMKLSSESKAKISE--TAGCGVADKAPKPEAVSLTELTAINPNARTEVNAOKIV 112
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 57 AILVAVDPEPIEKVTVG--SENAPEL--PETMTQAVESGAEPOVVGSVL 107
RESULT 12
AAG93285
ID AAG93285 standard; Protein: 579 AA.
```

```
AC AAG93285;
XX
XX
DT 13-SEP-2001 (first entry)
XX
XX
DE Human protein HP10637.
XX
XX
KW Human; gene therapy; tumour.
XX
XX
OS Homo sapiens.
XX
PN WO200142302-A1.
XX
PD 14-JUN-2001.
XX
PF 06-DEC-2000; 2000WO-JP08631.
XX
PR 06-DEC-1999; 99JP-0346863.
PR 06-DEC-1999; 99JP-0346864.
PR 08-FEB-2000; 2000JP-0031062.
PR 10-FEB-2000; 2000JP-0034090.
PR 10-FEB-2000; 2000JP-0034091.
PR 14-FEB-2000; 2000JP-0035829.
PR 14-FEB-2000; 2000JP-0035899.
PR 14-MAR-2000; 2000JP-0071161.
PR 30-MAY-2000; 2000JP-0160851.
XX
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
XX
PI Kato S, Eguchi C, Saeki M;
XX
XX
DR WPI: 2001-381646/40.
DR N-PSDB: AAH68570.
XX
XX
PT Human protein originated from tumor cell line, applicable as drug,
PT reagent for studying intracellular protein networks and protein source
PT for drug screening, also encoded cDNA for gene diagnosis and gene
PT therapy
XX
XX
PS Claim 1; Pages 275-278; 471pp; Japanese.
XX
XX
CC The present sequence is a human protein. The human protein, preferably
CC originated from tumor cell line, is applicable as a drug, a reagent for
CC studying intracellular protein networks and a protein source for
CC screening proteins for binding low molecular weight drugs. The human
CC protein coding sequence is useful for gene diagnosis and gene therapy,
CC expression vectors and transformant cells for detection of ligands and
CC receptors.
XX
SQ Sequence 579 AA;
Query Match 11.5%; Score 78; DB 22; Length 579;
Best Local Similarity 25.5%; Pred. No. 2.6;
Matches 25; Conservative 24; Mismatches 45; Indels 4; Gaps 1;
OY 18 ASCASTPESNPKNSSANLTSLIKHAVKOTCOTQLTGHOYKIAAMKLSSESKAKISETA 77
:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 323 SEASVSSSVAAKNSSSGSLTLPKSSSTNTSLTSKSTSOVAASLASKSSQTS--- 379
OY 78 CGCVADKAPKPEAVSLTELTAINPNARTEVNAOKIVRHS 115
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 380 -GSLVSKSTSLASVQLASKSSSQTSQSLPSKSTSQS 416
RESULT 13
ABB97409
ID ABB97409 standard; Protein: 560 AA.
XX
XX
AC ABB97409;
XX
XX
DT 27-JUN-2002 (first entry)
XX
XX
DE Novel human protein SEQ ID NO: 677.
```

KM Human; antianemic; vulnerary; antiinflammatory; immunomodulator;
 KM antileukemia; cerebroprotective; cytostatic; rheumatic; gene therapy;
 KM neutroprotective; antiparkinsonian; protein therapy; Est;
 KM expressed sequence tag.
 OS Homo sapiens.
 XX WO200222660-A2.
 XX 21-MAR-2002.
 XX 10-SEP-2001; 2001WO-US26015.
 XX 11-SEP-2000; 2000US-0659671.
 XX (HYSE-) HYSEQ INC.
 XX Tang YF, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX WPI: 2002-292408/33.
 DR N-PSDB: ABN32595.
 XX An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis -
 PS Claim 20; SEQ ID NO 677; 509pp; English.
 XX The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC hematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
 CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
 CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention.
 SQ Sequence 580 AA;
 Query Match 11.5%; Score 78; DB 23; Length 580;
 Best Local Similarity 25.5%; Pred. No. 2.6;
 Matches 25; Conservative 24; Mismatches 45; Indels 4; Gaps 1;
 QY 18 ASCATPESNPKNSANLITSLIKHAVKQTCOTQLGHOYWKIAAMKLSSEKAKISETA 77
 DB 324 SSEASVSSSVAKNSSSGTSLTPKSSSTNTSLTSKTSQVAASLASKSSQTS--- 380
 QY 78 CGCVADKAPPAVSLTETLTAAINPNAETVAOKIVRHS 115
 DB 381 -GSLVSKSTSLASVSQIASKSSSQTSQSLPKSTSGS 417
 RESULT 14
 ABB05596
 ID ABB05596 standard; Protein; 764 AA.
 XX ABB05596;
 XX 23-APR-2002 (first entry)
 DE Human testis development protein (PRTD).
 XX Human; testis development protein; PRTD; gene therapy.
 OS Homo sapiens.
 XX CN1318555-A.
 XX 24-OCT-2001.
 XX 11-APR-2001; 2001CN-0113501.

XX 11-APR-2001; 2001CN-0113501.
 PR (UYNA-) UNIV NANJING MEDICAL.
 XX Sha J, Zhou Z, Li J;
 PI WPI: 2002-115092/16.
 DR N-PSDB: ABA93410.
 XX Human testicular development relative protein gene encoded protein,
 PT useful for gene therapy -
 PS Claim 1; Page 1 (Claims); 7pp; Chinese.
 XX The present sequence represents the human testis development protein
 CC designated PRTD. The PRTD gene has a cDNA sequence of 2295 base pairs
 CC (bp) containing an open reading frame sequence of 764 bp from position
 CC 297 to 2591, having a Genbank number of AF111326. The present invention
 CC describes: (1) utilising the PRTD gene to prepare a fusion protein;
 CC (2) utilising the protein to immunise an animal and to prepare monoclonal
 CC and polyclonal antibodies; and (3) utilising the PRTD gene in preparing
 CC a testis development gene expressing chip. The expressed protein may be
 CC useful in gene therapy for treating related diseases.
 SQ Sequence 764 AA;
 Query Match 11.5%; Score 77.5; DB 23; Length 764;
 Best Local Similarity 22.3%; Pred. No. 4.5;
 Matches 40; Conservative 25; Mismatches 51; Indels 63; Gaps 7;
 QY 17 LASCATPESNPKNSANLITTS-----LIKHAVKQ-----COTQLG----- 54
 DB 418 LSSVSSPTSSPKTKVTVTSAQKSSQIGSSQLKRNHQRTEAVLTHQAOYPISSSEPPE 477
 QY 55 -----HOYWKIAAMKLSSEKAKIS-----ETACGCVADK---APE----- 87
 DB 478 EGEKEDLVOLKRRHPPSSPLGSKISKIRPKIKVSLISGDTAGGSCAPSGCAREGCKPI 537
 QY 88 -----AVSLTETLTAAINPNAETVA-----OKIVRHSILPKMLETYNATVPTT 132
 DB 538 TMTLGASAGAKELTGLTTAKSSSSGEGVASPVPVSSSTAPSAHLTLQSLRVATVS 596
 RESULT 15
 AAB95542
 ID AAB95542 standard; Protein; 789 AA.
 XX AAB95542;
 XX 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:18155.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 OS Homo sapiens.
 XX EP1074617-A2.
 XX 07-FEB-2001.
 XX 28-JUL-2000; 2000EP-0116126.
 XX 29-JUL-1999; 99JP-0248036.
 XX 27-AUG-1999; 99JP-0300253.
 XX 11-JAN-2000; 2000JP-0118776.
 XX 02-MAY-2000; 2000JP-0183767.
 XX 09-JUN-2000; 2000JP-0241899.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;


```
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-483447/52.
XX
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver.
XX
PS Claim 27; SEQ ID NO 27643; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 386 AA;
XX
Query Match 11.3%; Score 76.5; DB 22; Length 386;
Best Local Similarity 27.0%; Pred. No. 2.2;
Matches 37; Conservative 18; Mismatches 59; Indels 23; Gaps 6;
QY 7 VTATLMTAFILASCASTPESNPKSSANLTTSLIKHAVKOTCOTLGHQYKIAAKLS 66
DB 125 VSATVPKNNTPSVITSPSTAPNTASKMTTA-SKTAFTSTI-TSLPTVFTTSKITAG 182
QY 67 SE-----SKAKISETACGCVADKAPKAVSLT--ELTTAINPNARTEVQAOKIVRHS 115
DB 183 SEIPTASTTDSATTATSTKASGTTVESAPSTAPPTPAETTTASVPTTSTTGSENGHH- 241
QY 116 LKPCMLETVNAFTVPTT 132
DB 242 -----TVSS--VPTT 249
XX
RESULT 18
ABB20422
ID ABB20422 standard; Protein; 386 AA.
XX
AC ABB20422;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #2421 encoded by probe for measuring heart cell gene expression.
XX
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
```

```
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488899/53.
XX
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 15; SEQ ID NO 22192; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 386 AA;
XX
Query Match 11.3%; Score 76.5; DB 22; Length 386;
Best Local Similarity 27.0%; Pred. No. 2.2;
Matches 37; Conservative 18; Mismatches 59; Indels 23; Gaps 6;
QY 7 VTATLMTAFILASCASTPESNPKSSANLTTSLIKHAVKOTCOTLGHQYKIAAKLS 66
DB 125 VSATVPKNNTPSVITSPSTAPNTASKMTTA-SKTAFTSTI-TSLPTVFTTSKITAG 182
QY 67 SE-----SKAKISETACGCVADKAPKAVSLT--ELTTAINPNARTEVQAOKIVRHS 115
DB 183 SEIPTASTTDSATTATSTKASGTTVESAPSTAPPTPAETTTASVPTTSTTGSENGHH- 241
QY 116 LKPCMLETVNAFTVPTT 132
DB 242 -----TVSS--VPTT 249
XX
RESULT 19
AAM55821
ID AAM55821 standard; Protein; 386 AA.
XX
AC AAM55821;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 27926.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
```

PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00667.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483446/52.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 PS Example 4; SEQ ID NO: 27926; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.
 CC
 SQ Sequence 386 AA;

Query Match 11.3%; Score 76.5; DB 22; Length 386;
 Best Local Similarity 27.0%; Pred. No. 2.2; Mismatches 59; Indels 23; Gaps 6;
 Matches 37; Conservative 18;
 OY 7 VTATLMTAFTLASCASPTSPESNPKNSSANLTTSLIKHAVKOTCOTOLGHOYKIAAMKLS 66
 DB 125 VSATVPKNNTPSVITSTPSTAPNTASKMTMTTA-SKATATSTI-TSLPTVFTTTSKITAG 182
 OY 67 SE-----SKAKISETACGCYADKAPRAVSLT--ELTTAAINPARKREVAOKIYRHS 115
 DB 183 SEIPTASTSDSATAIATSKASGTTVESAPSTAPPTAETTTASVPTTSTTGSENGH- 241
 OY 116 LKPCMLETVNAFIVPTT 132
 DB 242 -----TVSS--VPTT 249

RESULT 20
 AAM68195
 ID AAM68195 standard; Protein; 386 AA.
 XX
 AC AAM68195;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 28501.
 XX
 KW Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 XX
 OS Homo sapiens.
 XX
 PN WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PR 30-JAN-2001; 2001WO-US00668.
 XX
 PF 04-FEB-2000; 2000US-0180312.
 XX
 PR 26-MAY-2000; 2000US-0207456.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.

PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX
 PS Example 4; SEQ ID NO: 28501; 658bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.
 CC
 SQ Sequence 386 AA;

Query Match 11.3%; Score 76.5; DB 22; Length 386;
 Best Local Similarity 27.0%; Pred. No. 2.2; Mismatches 59; Indels 23; Gaps 6;
 Matches 37; Conservative 18;
 OY 7 VTATLMTAFTLASCASPTSPESNPKNSSANLTTSLIKHAVKOTCOTOLGHOYKIAAMKLS 66
 DB 125 VSATVPKNNTPSVITSTPSTAPNTASKMTMTTA-SKATATSTI-TSLPTVFTTTSKITAG 182
 OY 67 SE-----SKAKISETACGCYADKAPRAVSLT--ELTTAAINPARKREVAOKIYRHS 115
 DB 183 SEIPTASTSDSATAIATSKASGTTVESAPSTAPPTAETTTASVPTTSTTGSENGH- 241
 OY 116 LKPCMLETVNAFIVPTT 132
 DB 242 -----TVSS--VPTT 249

RESULT 21
 AAM16015
 ID AAM16015 standard; Protein; 386 AA.
 XX
 AC AAM16015;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Peptide #2449 encoded by probe for measuring cervical gene expression.
 XX
 KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157278-A2.
 XX
 PD 09-AUG-2001.
 XX
 PR 30-JAN-2001; 2001WO-US00670.
 XX
 PF 04-FEB-2000; 2000US-0180312.
 XX
 PR 26-MAY-2000; 2000US-0207456.
 XX
 PR 30-JUN-2000; 2000US-0608408.
 XX
 PR 03-AUG-2000; 2000US-0632366.
 XX
 PR 21-SEP-2000; 2000US-0234687.
 XX
 PR 27-SEP-2000; 2000US-0236359.
 XX
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.

```

XX
DR WPI: 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
PS
XX
XX
Claim 27: SEQ ID NO 28785; 654PP; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see AAI31315-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
SQ Sequence 386 AA:
XX
Query Match 11.3%; Score 76.5; DB 22; Length 386;
Best Local Similarity 27.0%; Pred. No. 2.2;
Matches 37; Conservative 18; Mismatches 59; Indels 23; Gaps
QY 7 VTATLTMTAFTLASCASTPESNPKNSSANLTSLIKRAVKOTQOTOLGHOYKIAMKLS 66
DB 125 VSATVPKNNTPSPVINSTPTSTAPNTASKMTTA-SKTAITSTI-TSLPTVFYTTSTKITAG 182
QY 67 SE-----SKATSEFAGCGVADKAPAVSLT--ELTTAINEPNTAEVAKIVRHS 115
DB 183 SEIPTASTTDSATYTAISTKASGTVESAPSTAPTPAETTTASVPTTSTTSENHGH- 241
QY 116 LKPCMLETVNAFIPTT 132
DB 242 -----TVSS--VPTT 249
XX
RESULT 23
AAM03749
ID AAM03749 standard; Protein: 386 AA.
XX
AC AAM03749;
XX
DT 09-OCT-2001 (first entry)
XX
DE Peptide #2431 encoded by probe for measuring breast gene expression.
XX
KW Probe: human; breast disease; breast cancer; development disorder;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US00661.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SC, Hanzel DK, Chen W, Rank DR;
XX
WPI: 2001-476286/51.
XX
DR
XX
XX Novel single exon nucleic acid probe used to measuring gene expression
XX in a human breast -
XX
XX Claim 27: SEQ ID NO 12489; 322PP; English.
XX

```


PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match	11.28;	Score 76;	DB 21;	Length 310;
Best Local Similarity	29.38;	Pred. No.	1.9;	
Matches	37	Score	15	

Best Local Similarity 29.38; Pred. No. 1.9;
Matches 27; Conservative 15; Mismatches 44; Indels 6; Gaps 3.

0Y 16 TLASCSTP-ESNPKNSSANLTSLI--KHAVKQTCTQTGLGHQYWKTAAMKLSSSEYAK 72
11: :|||: . . .

DD 206 JMGACASKPRES DIVEGSVSTENAVVESKNATETDATTLOEKKESIEETKKEGETKED 265

QY 73 ISETACGCVADKAPFAVSLTELTAAINPNAR 104

Db 266 SSEAT--KAEPTPEAVKAEKTSSETEPPAQ 294

RESULT 25

ID AAG31396 standard; Protein; 357 AA.

AC AAG31396

DT 17-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 376996.

Protein Identification; signal transduction pathway; metabolic pathway;

termination sequence.

05 Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 09-MAR-1999; 99US-0123548

PR 25-MAR-1999; 99US-0126264

PR 01-APR-1999: 99US-0127462

08-APR-1999: 99TIS-0128714

19-APR-1999: 99US-0129845.

23-APR-1999; 99US-0130449.

23-APR-1999; 99US-0130891.

30-APR-1999; 99US-0132048.

04-MAY-1999; 99US-0132484.

06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137232.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139839.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142927.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144804.
PR 20-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

R 11-MAY-1999; 99US-0134256.

PR	14-MAY-1999;	9905-0134218
PR	14-MAY-1999;	9905-0134221
PR	14-MAY-1999;	9905-0134221
PR	14-MAY-1999;	9905-0134707
PR	18-MAY-1999;	9905-0134768
PR	19-MAY-1999;	9905-0134841
PR	20-MAY-1999;	9905-0135124
PR	21-MAY-1999;	9905-0135553
PR	24-MAY-1999;	9905-0136529
PR	25-MAY-1999;	9905-0136621
PR	27-MAY-1999;	9905-0136392
PR	28-MAY-1999;	9905-0136782
PR	01-JUN-1999;	9905-0137722
PR	03-JUN-1999;	9905-0137528
PR	04-JUN-1999;	9905-0137502
PR	07-JUN-1999;	9905-0137724
PR	08-JUN-1999;	9905-0138094
PR	10-JUN-1999;	9905-0138540
PR	10-JUN-1999;	9905-0138847
PR	14-JUN-1999;	9905-0139119
PR	16-JUN-1999;	9905-0139452
PR	16-JUN-1999;	9905-0139453
PR	16-JUN-1999;	9905-0139453
PR	18-JUN-1999;	9905-0139458
PR	18-JUN-1999;	9905-0139458
PR	18-JUN-1999;	9905-0139460
PR	18-JUN-1999;	9905-0139461
PR	18-JUN-1999;	9905-0139462
PR	18-JUN-1999;	9905-0139750
PR	18-JUN-1999;	9905-0139763
PR	21-JUN-1999;	9905-0139817
PR	22-JUN-1999;	9905-0139899
PR	23-JUN-1999;	9905-0140034
PR	23-JUN-1999;	9905-0140354
PR	24-JUN-1999;	9905-0140695
PR	28-JUN-1999;	9905-0140823
PR	29-JUN-1999;	9905-0140951
PR	30-JUN-1999;	9905-0141287
PR	01-JUL-1999;	9905-0141842
PR	02-JUL-1999;	9905-0142154
PR	02-JUL-1999;	9905-0142055
PR	06-JUL-1999;	9905-0142380
PR	08-JUL-1999;	9905-0142803
PR	09-JUL-1999;	9905-0142920
PR	12-JUL-1999;	9905-0142977
PR	13-JUL-1999;	9905-0143542
PR	14-JUL-1999;	9905-0143624
PR	15-JUL-1999;	9905-0144005
PR	16-JUL-1999;	9905-0144005
PR	16-JUL-1999;	9905-0144086
PR	19-JUL-1999;	9905-0144325
PR	19-JUL-1999;	9905-0144331
PR	19-JUL-1999;	9905-0144332
PR	19-JUL-1999;	9905-0144333
PR	19-JUL-1999;	9905-0144334
PR	19-JUL-1999;	9905-0144334
PR	21-JUL-1999;	9905-0144814
PR	21-JUL-1999;	9905-0145086
PR	21-JUL-1999;	9905-0145088
PR	22-JUL-1999;	9905-0145087
PR	22-JUL-1999;	9905-0145087
PR	22-JUL-1999;	9905-0145192
PR	23-JUL-1999;	9905-0145158
PR	23-JUL-1999;	9905-0145218

PR	23-JUL-1999;	9905-0145522;
PR	26-JUL-1999;	9905-0145526;
PR	27-JUL-1999;	9905-0145518;
PR	27-JUL-1999;	9905-0145518;
PR	27-JUL-1999;	9905-0145518;
PR	28-JUL-1999;	9905-0145591;
PR	02-AUG-1999;	9905-0146386;
PR	02-AUG-1999;	9905-0146388;
PR	02-AUG-1999;	9905-0146389;
PR	02-AUG-1999;	9905-0147038;
PR	04-AUG-1999;	9905-0147204;
PR	04-AUG-1999;	9905-0147302;
PR	05-AUG-1999;	9905-0147192;
PR	05-AUG-1999;	9905-0147260;
PR	06-AUG-1999;	9905-0147303;
PR	06-AUG-1999;	9905-0147493;
PR	09-AUG-1999;	9905-0147935;
PR	09-AUG-1999;	9905-0148171;
PR	10-AUG-1999;	9905-0148317;
PR	11-AUG-1999;	9905-0148341;
PR	12-AUG-1999;	9905-0148341;
PR	13-AUG-1999;	9905-0148565;
PR	13-AUG-1999;	9905-0148688;
PR	16-AUG-1999;	9905-0149368;
PR	17-AUG-1999;	9905-0149175;
PR	18-AUG-1999;	9905-0149426;
PR	20-AUG-1999;	9905-0149722;
PR	20-AUG-1999;	9905-0149929;
PR	22-AUG-1999;	9905-0149302;
PR	23-AUG-1999;	9905-0149302;
PR	23-AUG-1999;	9905-0149300;
PR	25-AUG-1999;	9905-0150566;
PR	26-AUG-1999;	9905-0150584;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151066;
PR	27-AUG-1999;	9905-0151080;
PR	30-AUG-1999;	9905-0151303;
PR	31-AUG-1999;	9905-0151438;
PR	01-SEP-1999;	9905-0151930;
PR	07-SEP-1999;	9905-0153363;
PR	10-SEP-1999;	9905-0153070;
PR	13-SEP-1999;	9905-0153758;
PR	16-SEP-1999;	9905-0154018;
PR	16-SEP-1999;	9905-0154039;
PR	20-SEP-1999;	9905-0155171;
PR	22-SEP-1999;	9905-0155139;
PR	23-SEP-1999;	9905-0155486;
PR	24-SEP-1999;	9905-0155659;
PR	28-SEP-1999;	9905-0156458;
PR	29-SEP-1999;	9905-0156596;
PR	04-OCT-1999;	9905-0157117;
PR	05-OCT-1999;	9905-0157573;
PR	06-OCT-1999;	9905-0157865;
PR	07-OCT-1999;	9905-0158025;
PR	08-OCT-1999;	9905-0158232;
PR	12-OCT-1999;	9905-0158331;
PR	13-OCT-1999;	9905-0159293;
PR	13-OCT-1999;	9905-0159294;
PR	14-OCT-1999;	9905-0159330;
PR	14-OCT-1999;	9905-0159330;
PR	14-OCT-1999;	9905-0159537;
PR	14-OCT-1999;	9905-0159637;
PR	18-OCT-1999;	9905-0159584;
PR	18-OCT-1999;	9905-0160741;
PR	21-OCT-1999;	9905-0160768;
PR	21-OCT-1999;	9905-0160768;
PR	21-OCT-1999;	9905-0160770;
PR	21-OCT-1999;	9905-0160814;
PR	21-OCT-1999;	9905-0160815;
PR	22-OCT-1999;	9905-0160981;

PR	22-OCT-1989;	9905-0160989;
PR	25-OCT-1989;	9905-0161404;
PR	25-OCT-1989;	9905-0161405;
PR	25-OCT-1989;	9905-0161406;
PR	26-OCT-1989;	9905-0161359;
PR	26-OCT-1989;	9905-0161360;
PR	26-OCT-1989;	9905-0161361;
PR	26-OCT-1989;	9905-0161920;
PR	28-OCT-1989;	9905-0161993;
PR	28-OCT-1989;	9905-0161992;
PR	29-OCT-1989;	9905-0162142;

Query Match	11.28;	Score 76;	DB 21;	Length 515;
Best Local Similarity	29.38;	Pred. NO. 3.9;		
Matches	27;	Conservative	15;	Mismatches 44;
			Indels	6;
			Gaps	3;

```

QY 16 TLASCASP--SSNNKNSANLITSLI--KHAVKOTCOTOLTHGYKMIAMKLSSEKAK 72
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 411 TMCAGASPKKSDIVEGSVSTENAVESKKNATITDNLTLQCKKEISIEETKKEGETKD 470
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 73 ISETACGCVADKAPEAVSLTELTAAINPNAR 104
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 471 SSEAT--KAEPTPEAVKAEKETSSETEPPAO 499
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 27
 AAY83275
 ID AAY83275 standard; Protein; 1203 AA

AC	AAV83275;
XX	
DT	16-AUG-2000 (first entry)

DE Candida albicans CSA1 surface protein.

KW Yeast: infection; detection; treatment; vaccine; antigen; CSa1;
probe; primer; vaginitis; immunoglobulin G; IgG.
KW

OS *Candida albicans*.

PN CA2274984-A1.

PD 10-JAN-2000.

PF 09-JUL-1999; 99CA-2274984.

PR 10-JUL-1998; 98CA-2237134.

PA (UYLA-) UNIV LAVAL.

PI Bourbonnais Y, Deslauriers N;

DR WPI; 2000-318497/28.

DR N-PSDB; AAZ93782.

PT Novel nucleic acid sequence encoding a *Candida albicans* surface antigen useful for detection, prophylaxis and treatment of candidiasis

PS Claim 5; Page 32-36; 50pp; English

CC The CSA1 surface antigen is unique to *Candida albicans* and so can
CC be exploited for the treatment of *Candida albicans* infections.
CC Nucleic acids, probes and primers specific to the CSA1 coding
CC sequence can be used to detect *C. albicans* infections in patients
CC by binding to *C. albicans* CSA1 DNA or mRNA in a patient sample. The
CC CSA1 protein, vaccines comprising it or antibodies directed against
CC it can be used to treat *C. albicans* infections. The virulence of
CC infection by *C. albicans* is reduced by binding of the antibody to
CC the surface antigen and masking it. Recurrent vaginitis in
CC immunocompetent patients can be treated by repeated administration
CC of CSA1 antigen to stimulate the production of immunoglobulin G
CC antibodies. The antibodies raised against CSA1 are specific for *C.*
CC *albicans* CSA1 antigen but not to other yeast or mycelial antigens,
CC

CC making detection and treatment easier.

XX Sequence 1203 AA;

Query Match 11.1%; Score 75; DB 21; Length 1203;
Best Local Similarity 24.5%; Pred. No. 17;
Matches 36; Conservative 25; Mismatches 54; Indels 32; Gaps 6;

OY 8 TATLMTAFTLASCSTPESNPKNSSANLTSTLKHAVK-----OTCOTQLT 53
DB 371 TSETSSOOLSSITSAPOSSATSSSTSTRTASINGFADKLYDQLEPCAKPCMFQNT 430
OY 54 G---HQYKIAMKISSSEKAKISFACGCVAD--KAPAVSLTELTTAAIN---PNART 105
DB 431 GITPCPYWDAGCLCYWPOFAGAIG---SCVADSKGQDIVSVTSLGTSVCSVAGVNAIPY 486
OY 106 EVAQKIVRHSLKPCMLETVNAFIPTT 132
DB 487 WMLPASVAKSSL-----SVAATAVPTTS 507

RESULT 28

AAy84814

ID AAY84814 standard; protein: 1203 AA.

XX AAY84814;

DT 08-AUG-2000 (first entry)

DE Amino acid sequence of the CSA1 gene of *Candida albicans*.

KW CSA1 gene; surface antigen; yeast; passive immunization;

KW *Candida albicans* infection; vaginal infection; systemic infection;

KW anti-fungal.

OS *Candida albicans*.

XX Key Location/Qualifiers

FT Peptide 1..17

FT Domain /note= "signal peptide" 42..143

FT Domain /note= "C-rich domain" 233..334

FT Domain /note= "C-rich domain" 403..504

FT Domain /note= "C-rich domain" 565..666

FT Domain /note= "C-rich domain" 750..851

FT Domain /note= "C-rich domain" 883

FT Modified-site /note= "putative N-glycosylation site" 893

FT Modified-site /note= "putative N-glycosylation site" 903

FT Modified-site /note= "putative N-glycosylation site" 913

FT Modified-site /note= "putative N-glycosylation site" 914

FT Modified-site /note= "putative N-glycosylation site" 927

FT Modified-site /note= "putative N-glycosylation site" 928

FT Modified-site /note= "putative N-glycosylation site" 937

FT Modified-site /note= "putative N-glycosylation site" 938

FT Modified-site /note= "putative N-glycosylation site" 954

FT Modified-site /note= "putative N-glycosylation site" 1184..1203

FT Peptide /note= "predicted GPI-anchoring determinant"

PN CA2237134-A1.

XX 10-JAN-2000.

PD 10-JUL-1998; 98CA-2237134.

PF 10-JUL-1998; 98CA-2237134.

PR 10-JUL-1998; 98CA-2237134.

XX (UYLA-) UNIV LAVAL.

XX Deslauriers N, Bourbonnais Y;

DR WPI; 2000-293503/26.

PT New CSA1 gene coding for a *Candida albicans* surface antigen, useful for

PT diagnosing the yeast form of *Candida albicans* within biological samples

PT by immunoassay or polymerase chain reaction -

PS Disclosure; Fig 4; 36pp; English.

CC The present sequence represents a protein encoded by the CSA1 gene.
CC The protein is a *Candida albicans* surface antigen. The CSA1 gene is
CC an intronless single copy gene in *Candida*, and RNA transcripts can
CC be detected in exponentially growing yeast cells. CSA1 gene expression
CC is strongly increased upon induction of the mycelial growth phase.
CC The CSA1 gene, its protein and antibodies against it are useful for
CC diagnosing the yeast form of *Candida albicans* within biological samples
CC by immunoassay or polymerase chain reaction (PCR). The antibody is
CC useful for passive immunization against *Candida albicans* infection.
CC The antibody is also useful in the treatment of vaginal or systemic
CC infection by *Candida*. The protein is useful as a therapeutic target
CC in the development of anti-fungal agents.

SQ Sequence 1203 AA;

Query Match 11.1%; Score 75; DB 21; Length 1203;
Best Local Similarity 24.5%; Pred. No. 17;
Matches 36; Conservative 25; Mismatches 54; Indels 32; Gaps 6;

OY 8 TATLMTAFTLASCSTPESNPKNSSANLTSTLKHAVK-----OTCOTQLT 53

DB 371 TSETSSOOLSSITSAPOSSATSSSTSTRTASINGFADKLYDQLEPCAKPCMFQNT 430

OY 54 G---HQYKIAMKISSSEKAKISFACGCVAD--KAPAVSLTELTTAAIN---PNART 105

DB 431 GITPCPYWDAGCLCYWPOFAGAIG---SCVADSKGQDIVSVTSLGTSVCSVAGVNAIPY 486

OY 106 EVAQKIVRHSLKPCMLETVNAFIPTT 132

DB 487 WMLPASVAKSSL-----SVAATAVPTTS 507

RESULT 29

AAR39286

ID AAR39286 standard; protein: 298 AA.

XX AAR39286;

DT 13-JAN-1994 (first entry)

DE Respiratory syncytial virus (RSV) G protein.

KW PIV; RSV; multimeric; hybrid; pathogen; chimeric protein; vaccine.

OS Respiratory syncytial virus.

PN MO9314207-A.

XX 22-JUL-1993.

PD 05-JAN-1993; 93MO-CA00001.

PF 06-JAN-1992; 92GB-0000117.

XX

Query Match 10.9%; Score 73.5; DB 23; Length 878;
Best Local Similarity 21.58; Pred. NO. 17;
Matches 39; Conservative 24; Mismatches 53; Indels 65; Gaps 7

RESULT 32	
AAU57395	
ID	AAU57395 standard; Protein; 89 AA
XX	

dermatological; osteopathic; neuroprotectant.

WO200181581-A2

01-NOV-2001

20-APR-2001; 2001WO-US12865.

21-APR-2000; 2000US-199047P.
03-TIN-2000; 2000US-200047P.

07-JUL-2000; 2000US-216747P.

A (CORI-) CORIXA CORP

N-PSDB; AAS59583.

PT Propionibacterium acnes polypeptides and nucleic acids useful for
vaccinating against and diagnosing infections, especially useful for
treating acne vulgaris -

ps Example 1; SEQ ID No 18590; 1069pp; English
xx

CC Sequences AAU39105-AAU68017 represent *Propionibacterium acnes* immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC *P. acnes*. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC *P. acnes* is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of *P. acnes* in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for *P. acnes* proteins. These antibodies can be used to
 CC downregulate expression and activity of *P. acnes* polypeptides and
 CC therefore treat *P. acnes* infections. The antibodies may also be used as
 CC diagnostic agents for determining *P. acnes* presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.

50 Sequence 89 AA

Query Match	10.8%;	Score 73;	DB 22;	Length 89;
Best Local Similarity	29.7%;	Pred. No. 0.72;		
Matches 19; Conservative	9;	Mismatches 32;	Indels 4;	Gaps 1.

QY 25 ESMPKSSSNLTLSLKIAVKTCTOLTGHQYWKIAAMLSSESKAKISETTACGVADK 8
 11 : : : : : : : : : : : : : : : :
 Db 6 ESALISSSVITTRIPRVAISTCSASAGSE---VANIKTLPGHARASTNCAPISTKN 6

QY	85	APEA	88
Db	62	SPDS	65

RESULT 33
AAU57478
ID AAU57478 standard; Protein; 197 AA
XX

AAU57478;

13-FEB-2002 (first entry)

Proionibacterium acnes immunogenic protein #18374

synovitis; synovitis; acne; pustulosis; hypertostis; osteomyelitis;
syndromes; syndromes; acne; pustulosis; hypertostis; osteomyelitis;

dermatological: osteonathic: neuroprotectant

Propionibacterium acnes

W0200181581-A2

01-NOV-2001

20-APR-2001; 2001WO-US12865.

21-APR-2000; 2000US-199047P.

Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

Arabidopsis thaliana.

EPI033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-0301439.

25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136788.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.

23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147419.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0148368.
PR 18-AUG-1999; 99US-0148175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.


```
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144864.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148655.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.

PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 10.7%; Score 72.5; DB 21; Length 436;
Best Local Similarity 26.5%; Pred. No. 8;
Matches 27; Conservative 19; Mismatches 43; Indels 13; Gaps 3;

QY 29 KNSANLTTSLIKH-AVKQTCQOTLGHQYWKIAMKLSSEKAK-----ISET 76
| : : : : : | : : : : : | : : : : : |
DB 332 KLSALEIESTLHPVAECVGLTNDYGEAVTALIIEMSAKKRREDESKVITLEE 391
| : : : : : | : : : : : | : : : : : |

QY 77 ACGGVADK-APEAVSLRELTAAINPARKTEVAKIVRHSK 117
| | | | | | : : : : : | : : : : : |
DB 392 LCGWAKDLAPYKLPTRLIWMESLPRNAMGVKNKEELKKSLE 433

RESULT 37
AAG18611
ID AAG18611 standard; Protein; 544 AA.
XX AAG18611;
AC AAG18611;
DE 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 20089.
XX Arabidopsis thaliana protein fragment SEQ ID NO: 20089.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
```

PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 23-APR-1999; 990S-0130891.
PR 28-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 07-MAY-1999; 990S-0132487.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139452.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140891.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.

PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148341.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0150566.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.

```

0Y      29  KNSANLTSLIKH-AVKOTCOTOLGHQYKIIAMKLSSEKAK-----ISET 76
      | : : : : : | : : : : : | : : : : : | : : : : : |
Db      440  KSALEFTESTLLEHPVAECVGLTDNDYGAVTAAIIIAESAKKRREDESKPVITLEE 499
      | | | | | | | | | | | | | | | | | | | | | |
0Y      77  ACGCVADK-APEAVSLTELTAAINDPNARTEVAOKIVRHSK 117
      | | | | | | | | | | | | | | | | | | | | | |
Db      500  LCGMAKDKLAPYKLPTRLIIWESLPRNMAKGVKKRELKSE 541
      | | | | | | | | | | | | | | | | | | | | | |

RESULT 38
AAC47314
ID      AAC47314 standard; Protein: 544 AA.
XX
XX      AAC47314;
AC
XX
DT      18-OCT-2000 (first entry)
DE
XX      Arabidopsis thaliana protein fragment SEQ ID NO: 59622.
DE
XX
XX      Protein identification: signal transduction pathway; metabolic pathway;
XX      hydralisation assay; genetic mapping; gene expression control; promoter;
XX      termination sequence.
XX
XX      Arabidopsis thaliana.
OS
XX      EPI033405-A2.
PN
XX
XX      06-SEP-2000.
PD

```

XX	25-FEB-2000;	2000EP-0301439.
XX		
PR	25-FEB-1999;	990S-0121825
PR	05-MAY-1999;	990S-0123180
PR	09-MAR-1999;	990S-0123548
PR	23-MAR-1999;	990S-0125788
PR	23-MAR-1999;	990S-0126264
PR	29-MAR-1999;	990S-0126785
PR	01-APR-1999;	990S-0127462
PR	06-APR-1999;	990S-0128234
PR	08-APR-1999;	990S-0128271
PR	16-APR-1999;	990S-0128845
PR	19-APR-1999;	990S-0130047
PR	21-APR-1999;	990S-0130449
PR	23-APR-1999;	990S-0130510
PR	28-APR-1999;	990S-0130891
PR	30-APR-1999;	990S-0131444
PR	30-APR-1999;	990S-0132048
PR	04-MAY-1999;	990S-0132407
PR	05-MAY-1999;	990S-0132484
PR	06-MAY-1999;	990S-0132485
PR	06-MAY-1999;	990S-0132486
PR	07-MAY-1999;	990S-0132487
PR	11-MAY-1999;	990S-0132863
PR	11-MAY-1999;	990S-0134256
PR	14-MAY-1999;	990S-0134218
PR	14-MAY-1999;	990S-0134219
PR	14-MAY-1999;	990S-0134221
PR	14-MAY-1999;	990S-0134370
PR	18-MAY-1999;	990S-0134768
PR	19-MAY-1999;	990S-0134941
PR	21-MAY-1999;	990S-0135124
PR	21-MAY-1999;	990S-0135353
PR	24-MAY-1999;	990S-0135629
PR	25-MAY-1999;	990S-0135829
PR	27-MAY-1999;	990S-0136021
PR	28-MAY-1999;	990S-0136302
PR	01-JUN-1999;	990S-0136762
PR	03-JUN-1999;	990S-0137232
PR	04-JUN-1999;	990S-0137528
PR	07-JUN-1999;	990S-0137724
PR	08-JUN-1999;	990S-0138050
PR	10-JUN-1999;	990S-0138504
PR	14-JUN-1999;	990S-0138847
PR	16-JUN-1999;	990S-0139119
PR	16-JUN-1999;	990S-0139452
PR	17-JUN-1999;	990S-0139453
PR	18-JUN-1999;	990S-0139492
PR	18-JUN-1999;	990S-0139454
PR	18-JUN-1999;	990S-0139455
PR	18-JUN-1999;	990S-0139456
PR	18-JUN-1999;	990S-0139457
PR	18-JUN-1999;	990S-0139458
PR	18-JUN-1999;	990S-0139460
PR	18-JUN-1999;	990S-0139461
PR	18-JUN-1999;	990S-0139462
PR	18-JUN-1999;	990S-0139463
PR	18-JUN-1999;	990S-0139750
PR	21-JUN-1999;	990S-0139763
PR	22-JUN-1999;	990S-0139817
PR	23-JUN-1999;	990S-0139899
PR	23-JUN-1999;	990S-0140353
PR	24-JUN-1999;	990S-0140354
PR	28-JUN-1999;	990S-0140695
PR	29-JUN-1999;	990S-0140823
PR	30-JUN-1999;	990S-0140991
PR	01-JUL-1999;	990S-0141287
PR	01-JUL-1999;	990S-0141842
PR	02-JUL-1999;	990S-0142154
PR	06-JUL-1999;	990S-0143055
PR	06-JUL-1999;	990S-0143390

PR	09-JUL-1999;	99US-0142803;
PR	09-JUL-1999;	99US-0142920;
PR	12-JUL-1999;	99US-0142977;
PR	13-JUL-1999;	99US-0143542;
PR	14-JUL-1999;	99US-0143624;
PR	15-JUL-1999;	99US-0144005;
PR	16-JUL-1999;	99US-0144086;
PR	19-JUL-1999;	99US-0144325;
PR	19-JUL-1999;	99US-0144331;
PR	19-JUL-1999;	99US-0144332;
PR	19-JUL-1999;	99US-0144333;
PR	19-JUL-1999;	99US-0144334;
PR	19-JUL-1999;	99US-0144335;
PR	19-JUL-1999;	99US-0144352;
PR	20-JUL-1999;	99US-0144362;
PR	20-JUL-1999;	99US-0144481;
PR	21-JUL-1999;	99US-0144814;
PR	21-JUL-1999;	99US-0145086;
PR	21-JUL-1999;	99US-0145088;
PR	22-JUL-1999;	99US-0145085;
PR	22-JUL-1999;	99US-0145087;
PR	22-JUL-1999;	99US-0145089;
PR	22-JUL-1999;	99US-0145143;
PR	23-JUL-1999;	99US-0145145;
PR	23-JUL-1999;	99US-0145212;
PR	23-JUL-1999;	99US-0145224;
PR	26-JUL-1999;	99US-0145276;
PR	27-JUL-1999;	99US-0145913;
PR	27-JUL-1999;	99US-0145918;
PR	28-JUL-1999;	99US-0145919;
PR	28-JUL-1999;	99US-0145951;
PR	02-AUG-1999;	99US-0146386;
PR	02-AUG-1999;	99US-0146388;
PR	02-AUG-1999;	99US-0147038;
PR	03-AUG-1999;	99US-0147038;
PR	03-AUG-1999;	99US-0147204;
PR	04-AUG-1999;	99US-0147302;
PR	05-AUG-1999;	99US-0147192;
PR	05-AUG-1999;	99US-0147260;
PR	06-AUG-1999;	99US-0147303;
PR	06-AUG-1999;	99US-0147416;
PR	09-AUG-1999;	99US-0147493;
PR	09-AUG-1999;	99US-0147935;
PR	10-AUG-1999;	99US-0148171;
PR	11-AUG-1999;	99US-0148319;
PR	12-AUG-1999;	99US-0148341;
PR	13-AUG-1999;	99US-0148565;
PR	13-AUG-1999;	99US-0148684;
PR	16-AUG-1999;	99US-0149368;
PR	17-AUG-1999;	99US-0149175;
PR	18-AUG-1999;	99US-0149426;
PR	20-AUG-1999;	99US-0149722;
PR	20-AUG-1999;	99US-0149723;
PR	23-AUG-1999;	99US-0149929;
PR	23-AUG-1999;	99US-0149902;
PR	25-AUG-1999;	99US-0149930;
PR	25-AUG-1999;	99US-0150566;
PR	26-AUG-1999;	99US-0150884;
PR	27-AUG-1999;	99US-0151065;
PR	27-AUG-1999;	99US-0151066;
PR	27-AUG-1999;	99US-0151080;
PR	30-AUG-1999;	99US-0151303;
PR	31-AUG-1999;	99US-0151438;
PR	01-SEP-1999;	99US-0151930;
PR	07-SEP-1999;	99US-0152363;
PR	10-SEP-1999;	99US-0153707;
PR	13-SEP-1999;	99US-0153758;
PR	15-SEP-1999;	99US-0154018;
PR	16-SEP-1999;	99US-0154039;
PR	20-SEP-1999;	99US-0154779;
PR	22-SEP-1999;	99US-0155139;
PR	23-SEP-1999;	99US-0155486;

PR	4-SEP-1999;	9905-01565655;
PR	28-SEP-1999;	9905-01565485;
PR	29-SEP-1999;	9905-01565596;
PR	04-OCT-1999;	9905-01571117;
PR	05-OCT-1999;	9905-01575753;
PR	06-OCT-1999;	9905-01578655;
PR	07-OCT-1999;	9905-01580029;
PR	08-OCT-1999;	9905-01582322;
PR	12-OCT-1999;	9905-01585369;
PR	13-OCT-1999;	9905-01592293;
PR	13-OCT-1999;	9905-01592294;
PR	13-OCT-1999;	9905-01595295;
PR	14-OCT-1999;	9905-01583329;
PR	14-OCT-1999;	9905-01553330;
PR	14-OCT-1999;	9905-01553331;
PR	14-OCT-1999;	9905-01596377;
PR	14-OCT-1999;	9905-01596378;
PR	18-OCT-1999;	9905-01595584;
PR	21-OCT-1999;	9905-01607411;
PR	21-OCT-1999;	9905-01607677;
PR	21-OCT-1999;	9905-01607688;
PR	21-OCT-1999;	9905-01607700;
PR	21-OCT-1999;	9905-01608615;
PR	21-OCT-1999;	9905-01608615;
PR	22-OCT-1999;	9905-01609080;
PR	22-OCT-1999;	9905-01609881;
PR	22-OCT-1999;	9905-01609881;
PR	25-OCT-1999;	9905-01614005;
PR	25-OCT-1999;	9905-01614005;
PR	25-OCT-1999;	9905-01614006;
PR	26-OCT-1999;	9905-01613559;
PR	26-OCT-1999;	9905-01613600;
PR	26-OCT-1999;	9905-01613611;
PR	28-OCT-1999;	9905-01619620;
PR	28-OCT-1999;	9905-01619620;
PR	29-OCT-1999;	9905-01619623;
PR	29-OCT-1999;	9905-01621492;

Query Match	10.7%	Score 72.5;	DB 21;	Length 544;
Best Local Similarity	26.5%	Pred. No. 11;		
Matches 27; Conservative	19;	Mismatches 43;	Indels 13;	Gaps 3;

```

0y 29 KNSNAUPTLSTLKH-AVKOTQOLOTGOWKIAMKLSSSSAK-----ISET 76
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 440 KLSAEIESTLHEPTYAACCVLTGDTNDGCAVTATIIAESAKKKREDESKPVITLEE 459
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0y 77 AGCGVADK-APEAVSLTELTAAINPAARFEVAKIVRSIK 117
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 500 LCGNAKDKIAYKLPTRLILNESIPRANAGVNNKELEKSSIE 541
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 39	
AA18610	
ID	AA18610 standard; Protein; 584 AA.
XX	
AC	AA18610;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 20088.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-0301439.
XX	

PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0128845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 28-APR-1999; 990S-0130891.
PR 30-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 04-MAY-1999; 990S-0132407.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 07-MAY-1999; 990S-0132487.
PR 11-MAY-1999; 990S-0132863.
PR 14-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 18-MAY-1999; 990S-0134370.
PR 19-MAY-1999; 990S-0134768.
PR 20-MAY-1999; 990S-0134941.
PR 21-MAY-1999; 990S-0135124.
PR 24-MAY-1999; 990S-0135353.
PR 25-MAY-1999; 990S-0135629.
PR 27-MAY-1999; 990S-0136021.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0141844.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.

PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144684.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145218.
PR 26-JUL-1999; 990S-0145224.
PR 27-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 03-AUG-1999; 990S-0146389.
PR 04-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 05-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148565.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149368.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 23-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 30-AUG-1999; 990S-0151080.
PR 31-AUG-1999; 990S-0151303.
PR 01-SEP-1999; 990S-0151430.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0153758.
PR 16-SEP-1999; 990S-0154018.
PR 20-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.


```
PR 16-JUL-1999; 99US-0144085;
PR 16-JUL-1999; 99US-0144086;
PR 19-JUL-1999; 99US-0144325;
PR 19-JUL-1999; 99US-0144326;
PR 19-JUL-1999; 99US-0144331;
PR 19-JUL-1999; 99US-0144332;
PR 19-JUL-1999; 99US-0144333;
PR 19-JUL-1999; 99US-0144334;
PR 20-JUL-1999; 99US-0144335;
PR 20-JUL-1999; 99US-0144336;
PR 20-JUL-1999; 99US-0144337;
PR 21-JUL-1999; 99US-0144884;
PR 21-JUL-1999; 99US-0144885;
PR 21-JUL-1999; 99US-0145086;
PR 22-JUL-1999; 99US-0145087;
PR 22-JUL-1999; 99US-0145088;
PR 22-JUL-1999; 99US-0145089;
PR 22-JUL-1999; 99US-0145090;
PR 23-JUL-1999; 99US-0145192;
PR 23-JUL-1999; 99US-0145193;
PR 23-JUL-1999; 99US-0145218;
PR 23-JUL-1999; 99US-0145219;
PR 26-JUL-1999; 99US-0145224;
PR 27-JUL-1999; 99US-0145276;
PR 27-JUL-1999; 99US-0145913;
PR 27-JUL-1999; 99US-0145918;
PR 28-JUL-1999; 99US-0145951;
PR 02-AUG-1999; 99US-0146386;
PR 02-AUG-1999; 99US-0146387;
PR 02-AUG-1999; 99US-0146388;
PR 03-AUG-1999; 99US-0146389;
PR 04-AUG-1999; 99US-0147038;
PR 04-AUG-1999; 99US-0147204;
PR 05-AUG-1999; 99US-0147302;
PR 05-AUG-1999; 99US-0147303;
PR 05-AUG-1999; 99US-0147192;
PR 06-AUG-1999; 99US-0147260;
PR 06-AUG-1999; 99US-0147303;
PR 09-AUG-1999; 99US-0147416;
PR 09-AUG-1999; 99US-0147493;
PR 10-AUG-1999; 99US-0147935;
PR 11-AUG-1999; 99US-0148171;
PR 12-AUG-1999; 99US-0148319;
PR 13-AUG-1999; 99US-0148341;
PR 13-AUG-1999; 99US-0148565;
PR 16-AUG-1999; 99US-0148684;
PR 17-AUG-1999; 99US-0149368;
PR 18-AUG-1999; 99US-0149175;
PR 20-AUG-1999; 99US-0149426;
PR 20-AUG-1999; 99US-0149722;
PR 20-AUG-1999; 99US-0149723;
PR 23-AUG-1999; 99US-0149829;
PR 23-AUG-1999; 99US-0149802;
PR 25-AUG-1999; 99US-0149930;
PR 26-AUG-1999; 99US-0150566;
PR 27-AUG-1999; 99US-0150884;
PR 27-AUG-1999; 99US-0151065;
PR 27-AUG-1999; 99US-0151066;
PR 30-AUG-1999; 99US-0151080;
PR 31-AUG-1999; 99US-0151303;
PR 01-SEP-1999; 99US-0151930;
PR 07-SEP-1999; 99US-0152363;
PR 10-SEP-1999; 99US-0153070;
PR 13-SEP-1999; 99US-0153070;
PR 15-SEP-1999; 99US-0153758;
PR 16-SEP-1999; 99US-0154018;
PR 16-SEP-1999; 99US-0154039;
PR 20-SEP-1999; 99US-0154779;
PR 22-SEP-1999; 99US-0155139;
PR 23-SEP-1999; 99US-0155486;
PR 24-SEP-1999; 99US-0155659;
PR 28-SEP-1999; 99US-0156458;
PR 29-SEP-1999; 99US-0156596;
PR 04-OCT-1999; 99US-0157117;
PR 05-OCT-1999; 99US-0157753;
PR 06-OCT-1999; 99US-0157865;
PR 07-OCT-1999; 99US-0158029;
PR 08-OCT-1999; 99US-0158232;
PR 12-OCT-1999; 99US-0158369;
PR 13-OCT-1999; 99US-0159293;
PR 13-OCT-1999; 99US-0159294;
PR 13-OCT-1999; 99US-0159295;
PR 14-OCT-1999; 99US-0159329;
PR 14-OCT-1999; 99US-0159330;
PR 14-OCT-1999; 99US-0159331;
PR 14-OCT-1999; 99US-0159637;
PR 18-OCT-1999; 99US-0159638;
PR 21-OCT-1999; 99US-0159584;
PR 21-OCT-1999; 99US-0160741;
PR 21-OCT-1999; 99US-0160767;
PR 21-OCT-1999; 99US-0160768;
PR 21-OCT-1999; 99US-0160770;
PR 21-OCT-1999; 99US-0160814;
PR 22-OCT-1999; 99US-0160815;
PR 22-OCT-1999; 99US-0160980;
PR 22-OCT-1999; 99US-0160981;
PR 22-OCT-1999; 99US-0160989;
PR 25-OCT-1999; 99US-0161404;
PR 25-OCT-1999; 99US-0161405;
PR 25-OCT-1999; 99US-0161406;
PR 26-OCT-1999; 99US-0161359;
PR 26-OCT-1999; 99US-0161360;
PR 26-OCT-1999; 99US-0161361;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161992;
PR 28-OCT-1999; 99US-0161993;
PR 29-OCT-1999; 99US-0162142;
PR 07-OCT-1999; 99US-0158029;
PR 08-OCT-1999; 99US-0158232;
PR 12-OCT-1999; 99US-0158369;
PR 13-OCT-1999; 99US-0159293;
PR 13-OCT-1999; 99US-0159294;
PR 13-OCT-1999; 99US-0159295;
PR 14-OCT-1999; 99US-0159329;
PR 14-OCT-1999; 99US-0159330;
PR 14-OCT-1999; 99US-0159331;
PR 14-OCT-1999; 99US-0159637;
PR 18-OCT-1999; 99US-0159638;
PR 21-OCT-1999; 99US-0159584;
PR 21-OCT-1999; 99US-0160741;
PR 21-OCT-1999; 99US-0160767;
PR 21-OCT-1999; 99US-0160768;
PR 21-OCT-1999; 99US-0160770;
PR 21-OCT-1999; 99US-0160814;
PR 22-OCT-1999; 99US-0160815;
PR 22-OCT-1999; 99US-0160980;
PR 22-OCT-1999; 99US-0160981;
PR 22-OCT-1999; 99US-0160989;
PR 25-OCT-1999; 99US-0161404;
PR 25-OCT-1999; 99US-0161405;
PR 25-OCT-1999; 99US-0161406;
PR 26-OCT-1999; 99US-0161359;
PR 26-OCT-1999; 99US-0161360;
PR 26-OCT-1999; 99US-0161361;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161992;
PR 28-OCT-1999; 99US-0161993;
PR 29-OCT-1999; 99US-0162142;
Query Match 10.7%; Score 72.5; DB 21; Length 584;
Best Local Similarity 26.5%; Pred. No. 12;
Matches 27; Conservative 19; Mismatches 43; Indels 13; Gaps 3;
OY 29 KSSANLTSILKH-AVKOTCOTLQGHQYKIAAKLSSSKAK-----1SET 76
| : : : : : | : : : : : | : : : : : | : : : : : |
Db 480 KLSALESTLEHPYAECCVGLTDNDGEAVTAIIASAAKKREDESKPVITLEE 539
OY 77 ACGCVADK-APEAVSLTELTTAAINPNARFEVAKIYRHSIK 117
| : : : : : | : : : : : | : : : : : | : : : : : |
Db 540 LCGWAKDKLAPYKLPULLIWESLPRNAMGVNKKELKKSLE 581
```

Search completed: April 28, 2003, 16:03:31
Job time : 43 secs

APPLICANT: EMASYSHN, Mary E
TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,963C
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/838,189
FILING DATE: 16-APR-1997
APPLICATION NUMBER: US 08/001,554
FILING DATE: 06-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 298 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-467-963C-8

Query Match 10.9%; Score 74; DB 2; Length 298;
Best Local Similarity 24.3%; Pred. No. 0.57;
Matches 27; Conservative 20; Mismatches 50; Indels 14; Gaps 3;

QY 1 MKILVYATATMTAFITASCSTPESNPKNSANLTTSLIKHA--VKOTCOTOLGHOY 57
DB 45 ILAMITSLITAIIFIASA-----NHKVTLTAIIDATSOIKNTPTPTLTDDPQ 96
QY 58 WKIAAMKLS---SESKAKISFTACGCVAADKAPKAVSLTETLTAIINPNT 105
DB 97 LGISFNSLSEITSOYTTTILASTTPGVKSNLOPTTVTKTKNTTQTQOPSKPT 147

RESULT 3
US-08-838-189D-8
Sequence 8, Application US/08838189D
Patent No. 5988169
GENERAL INFORMATION:
APPLICANT: KLEIN, Michel H
APPLICANT: DU, Run-Pan
APPLICANT: EMASYSHN, Mary E
TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto

STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/838,189D
FILING DATE: 16-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/001,554
FILING DATE: 06-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-687 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 298 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-838-189D-8

Query Match 10.9%; Score 74; DB 2; Length 298;
Best Local Similarity 24.3%; Pred. No. 0.57;
Matches 27; Conservative 20; Mismatches 50; Indels 14; Gaps 3;

QY 1 MKILVYATATMTAFITASCSTPESNPKNSANLTTSLIKHA--VKOTCOTOLGHOY 57
DB 45 ILAMITSLITAIIFIASA-----NHKVTLTAIIDATSOIKNTPTPTLTDDPQ 96
QY 58 WKIAAMKLS---SESKAKISFTACGCVAADKAPKAVSLTETLTAIINPNT 105
DB 97 LGISFNSLSEITSOYTTTILASTTPGVKSNLOPTTVTKTKNTTQTQOPSKPT 147

RESULT 4
US-08-852-344D-8
Sequence 8, Application US/08852344D
Patent No. 6017539
GENERAL INFORMATION:
APPLICANT: KLEIN, Michel H
APPLICANT: DU, Run-Pan
APPLICANT: EMASYSHN, Mary E
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION
TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,344D
FILING DATE: 07-MAY-1997

```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,639
; FILING DATE: 14-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-688 MTS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-852-344D-8

Query Match          10.9%; Score 74; DB 3; Length 298;
Best Local Similarity 24.3%; Pred. No. 0.57;
Matches 27; Conservative 20; Mismatches 50; Indels 14; Gaps 3;

QY 1 MKILYVATLMTAFLLASCASPESNPKNSSANLITSLIKHA---VKOTCOTOLTGHOY 57
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 45 ILAMISTSLITAIIFLISA-----NHKVTLTAIIDATSQIKNTPTVLTQDPQ 96
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 58 WKIAMKLS---SESKAKISETACGCVADKAPKAEVSLTETLTAAINPART 105
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 97 LGISFSNLSITQTTILASTPVGKSNLQPTVTKTKNTTTOPTOPSKPT 147
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 5
US-08-344-639E-8
; Sequence 8, Application US/08344639E
; Patent No. 6033668
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS
; TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS
; TITLE OF INVENTION: AND RESPIRATORY SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: 330 University Avenue, 6th floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/344,639E
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I

```

```

; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-391 MTS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TEXT: 065-24567 STMBAS
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
US-08-344-639E-8

Query Match          10.9%; Score 74; DB 3; Length 298;
Best Local Similarity 24.3%; Pred. No. 0.57;
Matches 27; Conservative 20; Mismatches 50; Indels 14; Gaps 3;

QY 1 MKILYVATLMTAFLLASCASPESNPKNSSANLITSLIKHA---VKOTCOTOLTGHOY 57
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 45 ILAMISTSLITAIIFLISA-----NHKVTLTAIIDATSQIKNTPTVLTQDPQ 96
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 58 WKIAMKLS---SESKAKISETACGCVADKAPKAEVSLTETLTAAINPART 105
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 97 LGISFSNLSITQTTILASTPVGKSNLQPTVTKTKNTTTOPTOPSKPT 147
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 6
US-08-467-969A-8
; Sequence 8, Application US/08467969A
; Patent No. 6168786
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; TITLE OF INVENTION: Chimeric Immunogens
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: 330 University Avenue, 6th floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,969A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-475 MTS:bh
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TEXT: 065-24567 STMBAS
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids

```

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-467-969A-8

Query Match 10.9%; Score 74; DB 4; Length 298;
Best Local Similarity 24.3%; Pred. No. 0.57;
Matches 27; Conservative 20; Mismatches 50; Indels 14; Gaps 3;

QY 1 MKKILVVTATMTAFILASCASPESNPKSSANLTTSLIKHA---VKOTCOTOLGHOY 57
DB 45 ILAMISTSLITAIIFISA-----NHRVTLTAIIOQDTSQIKNTPTLYLTQDPQ 96
QY 58 WKIAAMKLS---SESKAKISFACGCVADKAPAVSLTELTAAINPNART 105
DB 97 LGISFSLSEISQTTTILASTTPGVKSNLOPTTVTKTKNTTTTQTOPSKPT 147

RESULT 7

US-08-467-961A-8
Sequence 8, Application US/08467961A
Patent No. 6171783
GENERAL INFORMATION:
APPLICANT: Klein, Michel H
APPLICANT: Du, Run-Pan
APPLICANT: Ewasysbyn, Mary E
TITLE OF INVENTION: Chimeric Immunogens
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: 330 University Avenue, 6th Floor
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,961A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael J
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-476 MIS:bjh
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 298 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-467-961A-8

Query Match 10.9%; Score 74; DB 4; Length 298;
Best Local Similarity 24.3%; Pred. No. 0.57;
Matches 27; Conservative 20; Mismatches 50; Indels 14; Gaps 3;

QY 1 MKKILVVTATMTAFILASCASPESNPKSSANLTTSLIKHA---VKOTCOTOLGHOY 57
DB 45 ILAMISTSLITAIIFISA-----NHRVTLTAIIOQDTSQIKNTPTLYLTQDPQ 96
QY 58 WKIAAMKLS---SESKAKISFACGCVADKAPAVSLTELTAAINPNART 105
DB 97 LGISFSLSEISQTTTILASTTPGVKSNLOPTTVTKTKNTTTTQTOPSKPT 147

RESULT 8

US-08-001-554A-8
Sequence 8, Application US/08001554A
Patent No. 6225091
GENERAL INFORMATION:
APPLICANT: Klein, Michel H
APPLICANT: Du, Run-Pan
APPLICANT: Ewasysbyn, Mary E
TITLE OF INVENTION: Chimeric Immunogens
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: 330 University Avenue, 6th Floor
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/001,554A
FILING DATE: 06-JAN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael J
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-286
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 298 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
US-08-001-554A-8

Query Match 10.9%; Score 74; DB 4; Length 298;
Best Local Similarity 24.3%; Pred. No. 0.57;
Matches 27; Conservative 20; Mismatches 50; Indels 14; Gaps 3;

QY 1 MKKILVVTATMTAFILASCASPESNPKSSANLTTSLIKHA---VKOTCOTOLGHOY 57
DB 45 ILAMISTSLITAIIFISA-----NHRVTLTAIIOQDTSQIKNTPTLYLTQDPQ 96
QY 58 WKIAAMKLS---SESKAKISFACGCVADKAPAVSLTELTAAINPNART 105
DB 97 LGISFSLSEISQTTTILASTTPGVKSNLOPTTVTKTKNTTTTQTOPSKPT 147

RESULT 9

5258502-2
Patent No. 5258502
APPLICANT: Kuranda, Michael J
TITLE OF INVENTION: IMMOBILIZATION AND PURIFICATION OF
FUSION PROTEINS USING CHITIN-BINDING ABILITY
NUMBER OF SEQUENCES: 8

FILE REFERENCE: 003300-486
CURRENT APPLICATION NUMBER: US/09/087, 277B
CURRENT FILING DATE: 1998-05-29
EARLIER APPLICATION NUMBER: PCT/SE96/01558
EARLIER FILING DATE: 1996-11-28
EARLIER APPLICATION NUMBER: SE 9504272-7
EARLIER FILING DATE: 1995-11-29
EARLIER APPLICATION NUMBER: SE 9601506-0
EARLIER FILING DATE: 1996-04-19
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 878
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Description of Unknown Organism:bell gene (branching enzyme II)
US-09-087-277-2

Query Match 9.8%; Score 66.5; DB 4; Length 878;
Best Local Similarity 24.3%; Pred. No. 24;
Matches 28; Conservative 21; Mismatches 43; Indels 23; Gaps 4;

OY 26 SNPKSSANLTLTIKHAVKOTCOTQLTGHQYWKIAAMKLSSESAKISSETAC----- 78
DB 24 SNGDRRNANXSVFLKXHSLSR-----KILAEKSSYNSSESRPSTVAASGVLP 71
OY 79 GCVADKAPAVSLTELT-TAAINPNAFTEVAQKIVRHSLSKPCMLETVNAFIPTT 132
DB 72 GTQSDSSSSSTQDFEFTETSPENSASTVDSSSTMEHARQ---IKTENDVEPSS 123

RESULT 14
US-09-658-499-2
Sequence 2, Application US/09658499
Patent No. 6469231
GENERAL INFORMATION:
APPLICANT: EK, BO
APPLICANT: KHOSNOODI, Jamshid
APPLICANT: LARSSON, Clas-Tomas
APPLICANT: LARSSON, Hakan
APPLICANT: RASK, Lars
TITLE OF INVENTION: STARCH BRANCHING ENZYME II OF POTATO
FILE REFERENCE: 003300-486
CURRENT APPLICATION NUMBER: US/09/658, 499
CURRENT FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 09/087, 277
PRIOR FILING DATE: 1998-05-29
PRIOR APPLICATION NUMBER: PCT/SE96/01558
PRIOR FILING DATE: 1996-11-28
PRIOR APPLICATION NUMBER: SE 9504272-7
PRIOR FILING DATE: 1995-11-29
PRIOR APPLICATION NUMBER: SE 9601506-0
PRIOR FILING DATE: 1996-04-19
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 878
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Description of Unknown Organism:bell gene (branching enzyme II)
US-09-658-499-2

Query Match 9.8%; Score 66.5; DB 4; Length 878;
Best Local Similarity 24.3%; Pred. No. 24;
Matches 28; Conservative 21; Mismatches 43; Indels 23; Gaps 4;
OY 26 SNPKSSANLTLTIKHAVKOTCOTQLTGHQYWKIAAMKLSSESAKISSETAC----- 78
DB 24 SNGDRRNANXSVFLKXHSLSR-----KILAEKSSYNSSESRPSTVAASGVLP 71

OY 79 GCVADKAPAVSLTELT-TAAINPNAFTEVAQKIVRHSLSKPCMLETVNAFIPTT 132
DB 72 GTQSDSSSSSTQDFEFTETSPENSASTVDSSSTMEHARQ---IKTENDVEPSS 123

RESULT 15
US-08-602-725-32
Sequence 32, Application US/08602725
Patent No. 5965710
GENERAL INFORMATION:
APPLICANT: BODMER, WALTER F
APPLICANT: DURBIN, HELGA
APPLICANT: SNARY, DAVID
APPLICANT: STEWART, LORNA MD
APPLICANT: YOUNG, SUSAN
APPLICANT: BATES, PAUL A
TITLE OF INVENTION: MONOCLONAL ANTIBODIES FOR USE IN
DIGNOSIS AND TREATMENT OF COLORECTAL CANCER
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESSES:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602, 725
FILING DATE: 02-FEB-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB94/01816
FILING DATE: 19-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9317423
FILING DATE: 21-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36663
REFERENCE/DOCKET NUMBER: 1090-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4091
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 464 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-602-725-32

Query Match 9.8%; Score 66; DB 2; Length 464;
Best Local Similarity 25.7%; Pred. No. 11;
Matches 29; Conservative 13; Mismatches 45; Indels 26; Gaps 4;
OY 26 SNPKSSANLTLTIKHAVKOTCOTQLTGHQY-WKIAAMKLSSESAKISSETACGCVADK 84
DB 326 AKPOIAKSKTYTGDKDSVNLTCSTNDTGISIRMFKNOSLPSSEEMKLSQGN----- 378
OY 85 APEAVSLTELTAAINPNR-----TEVAQIVRHSLSKPCMLETVNAFIYP 130
DB 379 -----TTLISNPVAKREDAGTYWCEVFNPIKNSQDPIML-NVYNALP 420

RESULT 16
US-08-740-223A-16
Sequence 16, Application US/08740223A
Patent No. 6265564

QY	128	IVPTTT	133
DB	466	TTTTTT	471
QY	20	RESULT	21
DB	US-09-081-320-3	US-09-574-141A-3	US-09-574-141A-3
QY	Sequence 3, Application US/09081320	Sequence 3, Application US/09574141A	Sequence 3, Application US/09574141A
DB	Patent No. 6093544	Patent No. 6093544	Patent No. 6093544
QY	GENERAL INFORMATION:	GENERAL INFORMATION:	GENERAL INFORMATION:
DB	APPLICANT: Gonsalves, Dennis	APPLICANT: Gonsalves, Dennis	APPLICANT: Gonsalves, Dennis
QY	APPLICANT: Meng, Baozhong	APPLICANT: Meng, Baozhong	APPLICANT: Meng, Baozhong
DB	TITLE OF INVENTION: RUPESTRIS STEM PITTING ASSOCIATED VIRUS	TITLE OF INVENTION: RUPESTRIS STEM PITTING ASSOCIATED VIRUS	TITLE OF INVENTION: RUPESTRIS STEM PITTING ASSOCIATED VIRUS
QY	TITLE OF INVENTION: NUCLEIC ACIDS, PROTEINS, AND THEIR USES	TITLE OF INVENTION: NUCLEIC ACIDS, PROTEINS, AND THEIR USES	TITLE OF INVENTION: NUCLEIC ACIDS, PROTEINS, AND THEIR USES
DB	NUMBER OF SEQUENCES: 54	NUMBER OF SEQUENCES: 54	NUMBER OF SEQUENCES: 54
QY	CORRESPONDENCE ADDRESS:	CORRESPONDENCE ADDRESS:	CORRESPONDENCE ADDRESS:
DB	ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP	ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP	ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
QY	STREET: Clinton Square, P.O. Box 1051	STREET: Clinton Square, P.O. Box 1051	STREET: Clinton Square, P.O. Box 1051
DB	CITY: Rochester	CITY: Rochester	CITY: Rochester
QY	STATE: New York	STATE: New York	STATE: New York
DB	COUNTRY: U.S.A.	COUNTRY: U.S.A.	COUNTRY: U.S.A.
QY	ZIP: 14603	ZIP: 14603	ZIP: 14603
DB	COMPUTER READABLE FORM:	COMPUTER READABLE FORM:	COMPUTER READABLE FORM:
QY	MEDIUM TYPE: Floppy disk	MEDIUM TYPE: Floppy disk	MEDIUM TYPE: Floppy disk
DB	COMPUTER: IBM PC compatible	COMPUTER: IBM PC compatible	COMPUTER: IBM PC compatible
QY	OPERATING SYSTEM: PC-DOS/MS-DOS	OPERATING SYSTEM: PC-DOS/MS-DOS	OPERATING SYSTEM: PC-DOS/MS-DOS
DB	SOFTWARE: PatentIn Release #1.0, Version #1.30	SOFTWARE: PatentIn Release #1.0, Version #1.30	SOFTWARE: PatentIn Release #1.0, Version #1.30
QY	CURRENT APPLICATION DATA:	CURRENT APPLICATION DATA:	CURRENT APPLICATION DATA:
DB	APPLICATION NUMBER: US/09/081,320	APPLICATION NUMBER: US/09/081,320	APPLICATION NUMBER: US/09/081,320
QY	FILING DATE:	FILING DATE:	FILING DATE:
DB	CLASSIFICATION:	CLASSIFICATION:	CLASSIFICATION:
QY	PRIOR APPLICATION DATA:	PRIOR APPLICATION DATA:	PRIOR APPLICATION DATA:
DB	APPLICATION NUMBER: US 60/047,147	APPLICATION NUMBER: US 60/047,147	APPLICATION NUMBER: US 60/047,147
QY	FILING DATE: 20-MAY-1997	FILING DATE: 20-MAY-1997	FILING DATE: 20-MAY-1997
DB	PRIOR APPLICATION DATA:	PRIOR APPLICATION DATA:	PRIOR APPLICATION DATA:
QY	APPLICATION NUMBER: US 60/069,902	APPLICATION NUMBER: US 60/069,902	APPLICATION NUMBER: US 60/069,902
DB	FILING DATE: 17-DEC-1997	FILING DATE: 17-DEC-1997	FILING DATE: 17-DEC-1997
QY	ATTORNEY/AGENT INFORMATION:	ATTORNEY/AGENT INFORMATION:	ATTORNEY/AGENT INFORMATION:
DB	NAME: Goldman, Michael L.	NAME: Goldman, Michael L.	NAME: Goldman, Michael L.
QY	REGISTRATION NUMBER: 30,727	REGISTRATION NUMBER: 30,727	REGISTRATION NUMBER: 30,727
DB	REFERENCE/DOCKET NUMBER: 19603/1722	REFERENCE/DOCKET NUMBER: 19603/1722	REFERENCE/DOCKET NUMBER: 19603/1722
QY	TELEPHONE: (716) 263-1304	TELEPHONE: (716) 263-1304	TELEPHONE: (716) 263-1304
DB	TELEFAX: (716) 263-1600	TELEFAX: (716) 263-1600	TELEFAX: (716) 263-1600
QY	INFORMATION FOR SEQ ID NO: 3:	INFORMATION FOR SEQ ID NO: 3:	INFORMATION FOR SEQ ID NO: 3:
DB	SEQUENCE CHARACTERISTICS:	SEQUENCE CHARACTERISTICS:	SEQUENCE CHARACTERISTICS:
QY	LENGTH: 2161 amino acids	LENGTH: 2161 amino acids	LENGTH: 2161 amino acids
DB	TYPE: amino acid	TYPE: amino acid	TYPE: amino acid
QY	STRANDEDNESS:	STRANDEDNESS:	STRANDEDNESS:
DB	TOPOLOGY: linear	TOPOLOGY: linear	TOPOLOGY: linear
QY	MOLECULE TYPE: protein	MOLECULE TYPE: protein	MOLECULE TYPE: protein
DB	US-09-081-320-3	US-09-081-320-3	US-09-081-320-3
QY	Query Match	Query Match	Query Match
DB	Best Local Similarity 9.6%; Score 65; DB 3; Length 2161;	Best Local Similarity 20.7%; Score 65; DB 3; Length 2161;	Best Local Similarity 20.7%; Score 65; DB 3; Length 2161;
QY	Matches 28; Conservative 19; Mismatches 58; Indels 30; Gaps 3;	Matches 28; Conservative 19; Mismatches 58; Indels 30; Gaps 3;	Matches 28; Conservative 19; Mismatches 58; Indels 30; Gaps 3;
DB	11 LMTATTLASCASTPSPNPKNSANLITSLIKHAVVQTCOTOL-----TGHQYWK 59	11 LMTATTLASCASTPSPNPKNSANLITSLIKHAVVQTCOTOL-----TGHQYWK 59	11 LMTATTLASCASTPSPNPKNSANLITSLIKHAVVQTCOTOL-----TGHQYWK 59
QY	60 IAAK-----LSSEKAKISEFAAGCVADKAPKPEAVSLTELTA-AIN 100	60 IAAK-----LSSEKAKISEFAAGCVADKAPKPEAVSLTELTA-AIN 100	60 IAAK-----LSSEKAKISEFAAGCVADKAPKPEAVSLTELTA-AIN 100
DB	564 VGLFKHIALTHCFMNSGCLQWFLRLQRBNSLFLKDKRASSFADLDCVEIKVYOLVTSQAIL 623	564 VGLFKHIALTHCFMNSGCLQWFLRLQRBNSLFLKDKRASSFADLDCVEIKVYOLVTSQAIL 623	564 VGLFKHIALTHCFMNSGCLQWFLRLQRBNSLFLKDKRASSFADLDCVEIKVYOLVTSQAIL 623
QY	101 PNAREVAOKIVRHS 115	101 PNAREVAOKIVRHS 115	101 PNAREVAOKIVRHS 115
DB	624 PEALLSLTKVFEVRS 638	624 PEALLSLTKVFEVRS 638	624 PEALLSLTKVFEVRS 638

Db 78 PSAIAAPAPSS-STYTTSTSTNPACVCKPTDSMSQRKRSKRTOHMKVI---IKPPSP 133
QY 119 -CMLE 122
111:
Db 134 TCMK 138

RESULT 24

US-08-465-388-93
; Sequence 93, Application US/08465388
; Patent No. 5753468

GENERAL INFORMATION:

APPLICANT: WILLIAM JOHN MARTIN
TITLE OF INVENTION: ISOLATED STEALTH VIRUSES
TITLE OF INVENTION: AND RELATED VACCINES
NUMBER OF SEQUENCES: 104

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,388
FILING DATE: June 5, 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA: Including application
PRIOR APPLICATION DATA: described below:

Four

APPLICATION NUMBER: 08/157,811
FILING DATE: No. 5753488ember 23, 1993
APPLICATION NUMBER: 07/887,502
FILING DATE: May 22, 1992
APPLICATION NUMBER: 07/704,814
FILING DATE: May 23, 1991
APPLICATION NUMBER: 07/763,039
FILING DATE: September 20, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 213/300
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 93:
SEQUENCE CHARACTERISTICS:
LENGTH: 264 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-465-388-93

Query Match

Best Local Similarity 9.5%; Score 64.5; DB 1; Length 264;
Matches 32; Conservative 17; Mismatches 37; Indels 39; Gaps 7;

QY 30 NSSANLTSLIKHAVKQTCOT-----QLNGHQWKIAW-----KLSSESKAKISTT 76
Db 21 SSSAGVTPAPSSMTTAPTAIPTTAIQVPGMO---ITSLGCTPKPKSKPKKIPAP 77
QY 77 ACGVADAPAPAVSLTETTAIINP-----NARTEVAOKIVRSILKP----- 118
Db 78 PSAIAAPAPSS-STYTTSTSTNPACVCKPTDSMSQRKRSKRTOHMKVI---IKPPSP 133

QY 119 -CMLE 122
111:
Db 134 TCMK 138

RESULT 25

US-09-413-814-42
; Sequence 42, Application US/09413814
; Patent No. 6225064

GENERAL INFORMATION:

APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Beyer, Stefan
APPLICANT: Bioecker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M

APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hofle, Gerhard
APPLICANT: Mueller, Joachim

APPLICANT: Reichenbach, Hans

TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or
FILE REFERENCE: PCT/US 99/23535

CURRENT APPLICATION NUMBER: US/09/413,814

CURRENT FILING DATE: 1999-10-07

EARLIER FILING DATE: 1998-10-09

NUMBER OF SEQ ID NOS: 107

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 42

LENGTH: 2539

TYPE: PRT

ORGANISM: Sorangium cellulosum

US-09-413-814-42

Query Match

Best Local Similarity 9.5%; Score 64.5; DB 4; Length 2539;
Matches 23; Conservative 15; Mismatches 39; Indels 9; Gaps 3;

QY 42 HAVKQTCOTQLTGHQWKIAWKLSSSEKAKISEFACGCVADKAPAVSLTETTAIINP 101
Db 750 HAVKGCATNNDGDR--KAGTYSVSAQQAQAVIRSA-QILDVAPESTIVEA-----H 800

QY 102 NARTEVAOKIVRSILKPCMLETVNAF 127
Db 801 GTGTKLDSIEIKALQAFASDKNGF 826

RESULT 26

US-08-936-165A-516

; Sequence 516, Application US/08936165A

; Patent No. 6348582

GENERAL INFORMATION:

APPLICANT: Black, Michael

APPLICANT: Burnham, Martin

APPLICANT: Hodgson, John

APPLICANT: Knowles, David

APPLICANT: Lonetto, Michael

APPLICANT: Pratt, Julie

APPLICANT: Reichard, Richard

APPLICANT: Rosenberg, Martin

APPLICANT: Ward, Judith

TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
TITLE OF INVENTION: Polypeptides and their uses

NUMBER OF SEQUENCES: 534

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA

```

: COUNTRY: USA
: ZIP: 19406-0939
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/936,165A
: FILING DATE: 24-SEP-1997
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/027,032
: FILING DATE: 24-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Glimi, Edward R
: REGISTRATION NUMBER: 38,891
: REFERENCE/DOCKET NUMBER: P50549
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-270-4478
: TELEFAX: 610-270-5090
: TELEX:
: INFORMATION FOR SEQ ID NO: 516:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 276 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: Protein
: US-08-936-165A-516

```

```

Query Match          9.5%; Score 64; DB 4; Length 276;
Best Local Similarity 25.2%; Pred. No. 8.7;
Matches 27; Conservative 16; Mismatches 44; Indels 20; Gaps 4;

```

```

OY 17 LASCASTPESPKSSANLTTSLIKHAVKQTCQTLGHQYWKIAAMKLSSSKA----- 71
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 106 IISAXNPEMP-----DIIQKASQVNSAKSALKGDE--KLAARQTAKSQIDGRVTD 156
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 72 --KISFTAGCGVADKAPKPEAVSLTELTAAINPNARTEVAOKIVRHSI 116
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 157 LNNQRITAXNAEVDQAPXLAIV---TAAKKATSLNTAMGNVXKHAL 199
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 27
US-08-209-521-11
: Sequence 11, Application US/08209521
: Patent No. 5922855
: GENERAL INFORMATION:
: APPLICANT: Liskay, Robert M.
: APPLICANT: Bronner, C. Eric
: APPLICANT: Baker, Sean M.
: APPLICANT: Bollag, Roni J.
: APPLICANT: Kolodner, Richard D.
: TITLE OF INVENTION: MAMMALIAN DNA MISMATCH REPAIR GENES
: NUMBER OF SEQUENCES: 30
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Kolisch, Hartwell, Dickinson, McCormack &
: STREET: 520 S.W. Yamhill, Suite 200
: CITY: Portland
: STATE: Oregon
: COUNTRY: US
: ZIP: 97204
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/209,521
: FILING DATE: 08-MAR-1994

```

```

: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Van Rysselberghe, Pierre C.
: REGISTRATION NUMBER: 33,557
: REFERENCE/DOCKET NUMBER: OHSU 306A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (503) 224-6655
: TELEFAX: (503) 295-6679
: INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 341 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: POSITION IN GENOME:
: MAP POSITION: 3p21.3-23
: US-08-209-521-11

```

```

Query Match          9.5%; Score 64; DB 2; Length 341;
Best Local Similarity 25.7%; Pred. No. 12;
Matches 26; Conservative 9; Mismatches 60; Indels 6; Gaps 1;

```

```

OY 8 TATLMTAFTLASCASTPESPKSSANLTTSLIKHAVKQTCQTLGHQYWKIAAMKLSS 67
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 232 TGTATTTCACACATAGATAAATAGTTG-----GTACTTTTACTTGTAAATGTA 285
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 68 ESKAKISFTAGCGVADKAPKPEAVSLTELTAAINPNARTEVA 108
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 286 TGCAAATGTATGCAAAATCTGHCACAACTTAATGADCTTTAA 326
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 28
US-08-078-683A-8
: Sequence 8, Application US/08078683A
: Patent No. 5486599
: GENERAL INFORMATION:
: APPLICANT: Saunders, Scott
: APPLICANT: Bernfield, Merton
: APPLICANT: Kato, Masato
: TITLE OF INVENTION: Construction and Use of Synthetic
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHYE & COCKFIELD
: STREET: 60 State Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII (text)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/078,683A
: FILING DATE: 17-JUN-1993
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Vincent, Matthew P.
: REGISTRATION NUMBER: 36,709
: REFERENCE/DOCKET NUMBER: CME-062
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 227-7400
: TELEFAX: (617) 227-5941
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 357 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: FRAGMENT TYPE: internal

```


APPLICANT: KRAUS, MATTHIAS H.; AARONSON, STUART A.
TITLE OF INVENTION: DNA SEGMENT ENCODING A GENE FOR A
RECEPTOR RELATED TO THE EPIDERMAL GROWTH FACTOR RECEPTOR
NUMBER OF SEQUENCES: 5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/444,406
FILING DATE: 01-DEC-1989
SEQ ID NO: 4
LENGTH: 1343
5183884-4

Query Match
Best Local Similarity 23.5%; Score 64; DB 6; Length 1343;
Matches 35; Conservative 23; Mismatches 49; Indels 42; Gaps 7;

Db 16 TLASCATPE---SNPKNSANLTTSS---LTKHAKVOTCOTQLTGHQYKIAAMKLS 67
1027 TLGSAALSLPVGLNLRPGSQSLSPSSGYMPMGNLSSGCE-----SAVSGSS 1076

Qy 68 E---SKAKISFACCCVADKADPEA-----VSLTELTTAIPNARTEVAOKIV 112
1077 ERCPRPVSLHMPRGCLASSESEGHVTSSEAFLOEKVSMCRSRSRSPRPAGDSAYHSQ 1136

Qy 113 RHS-----LKPCL--ETVNAFTVPTT 132
1137 RLSLTPVTPPLSPGLEEDVNGVMPDT 1165

RESULT 35
US-08-602-713-12
Sequence 12, Application US/08602713
Patent No. 5798205

GENERAL INFORMATION:
APPLICANT: Hauser, Hans-Peter; Knapp Stefan; Brust, Stefan;
APPLICANT: G rler, Lutz G.; Eberle, Josef; Kaplue, Lazare;
APPLICANT: Zekeng, L opold Achengul
TITLE OF INVENTION: Retrovirus from The HIV Group And Its Use
TITLE OF INVENTION: (MVP-2901/94)
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felife & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,713
FILING DATE: 16-FEBRUARY-1996
CLASSIFICATION: 436

PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 195 05 262
FILING DATE: 16-FEBRUARY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5798205man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LEDER 203
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 356 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: Protein
FRAGMENT TYPE: Internal

US-08-602-713-12

Query Match
Best Local Similarity 9.4%; Score 63.5; DB 1; Length 356;
Matches 26; Conservative 13; Mismatches 47; Indels 7; Gaps 2;

Qy 28 PKNSANLTTSLIKHAKVOTCOTQLTGHQYKIAAMKLSSESAKISFACCGVADKADPE 87
Db 185 PYNKSSNITFRPIGDMKIDIMRTOMYN---YKVRVKSFSVAPTAKISRPIVIGTNNQREKR 241

Qy 88 AVSLTEL---TTAIPNARTEVAOKIVRHS 116
Db 242 AVGLGMLFLGVLSAAGSTWCAAGVTLSVTRHSL 274

RESULT 36
US-08-989-493-12
Sequence 12, Application US/08989493
Patent No. 6162631

GENERAL INFORMATION:
APPLICANT: Hauser, Hans-Peter; Knapp Stefan; Brust, Stefan;
APPLICANT: G rler, Lutz G.; Eberle, Josef; Kaplue, Lazare;
APPLICANT: Zekeng, L opold Achengul
TITLE OF INVENTION: Retrovirus from The HIV Group And Its Use
TITLE OF INVENTION: (MVP-2901/94)
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felife & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/989,493
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/602,713
FILING DATE: 16-FEBRUARY-1996
APPLICATION NUMBER: DE 195 05 262
FILING DATE: 16-FEBRUARY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6162631man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LEDER 203
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 356 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: Protein
FRAGMENT TYPE: Internal

US-08-989-493-12

Query Match
Best Local Similarity 9.4%; Score 63.5; DB 4; Length 356;
Matches 26; Conservative 13; Mismatches 47; Indels 7; Gaps 2;

Qy 28 PKNSANLTTSLIKHAKVOTCOTQLTGHQYKIAAMKLSSESAKISFACCGVADKADPE 87
Db 185 PYNKSSNITFRPIGDMKIDIMRTOMYN---YKVRVKSFSVAPTAKISRPIVIGTNNQREKR 241

Qy 88 AVSLTEL---TTAIPNARTEVAOKIVRHS 116

Db 242 AVGLGMLFLCVLSAAGSTMGAGATVLSVTRHSL 274

RESULT 37

US-09-134-001C-4463
Sequence 4463, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4463
LENGTH: 2137
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4463

Query Match
Best Local Similarity 21.8%; Pred. No. 2.1e+02;
Matches 27; Conservative 24; Mismatches 50; Indels 23; Gaps 4;

QY 8 TATLTATFLASCA-----STPESNPKNSANLTTSL-----IKHAVKOTQOTLTGH 55
Db 1091 TSTLSGTSASTSDASTSTSESDTSLSTSVSDTSASSTSESASTS-TSE 1149
QY 56 QYKIAAMKLSSESKAKISERACGCVADKA-----PAVSLTETLTAINPNART 105
Db 1150 SESNASYSLSGSLSTSDSTSTSTSDASTSTSESDTSLSTSESTSLSDSTST 1209
QY 106 EVAQ 109
Db 1210 STSE 1213

RESULT 38

US-08-732-228-4
Sequence 4, Application US/08732228
Patent No. 5981469
GENERAL INFORMATION:
APPLICANT: ANDERSSON, Mats
APPLICANT: BOMAN, Hans G.
APPLICANT: JORNVAL, Hans
APPLICANT: MUTT, Viktor
TITLE OF INVENTION: NEW 78 RESIDUE POLYPEPTIDE (NK-LYSINE)
TITLE OF INVENTION: AND ITS USE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/732,228
FILING DATE: 04-DEC-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/SE95/00475

FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9401480-0
FILING DATE: 29-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 003300-388
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 145 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-732-228-4

Query Match
Best Local Similarity 27.4%; Score 63; DB 2; Length 145;
Matches 20; Conservative 13; Mismatches 30; Indels 10; Gaps 3;

QY 56 QYKIAAMKLSSESKAKISERACGCVADKAPAVSLTETLTAINPNART--EVAQKIYR 113
Db 27 EYDIDARAHLEBEK-----SCPCLAGSGPGQDLTK--TQELGRDYRTCLTIYOKLAK 78
QY 114 HSLKPCMEETVNA 126
Db 79 MYDKPTQRSVNA 91

RESULT 39
5169835-8
Patent No. 5169835
APPLICANT: WAI-YEE, CHAN
TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS
NUMBER OF SEQUENCES: 48
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/390,409
FILING DATE: 07-AUG-1989
SEQ ID NO: 8
LENGTH: 144
5169835-8

Query Match
Best Local Similarity 31.7%; Pred. No. 5;
Matches 20; Conservative 5; Mismatches 35; Indels 3; Gaps 2;

QY 15 FTLASCSTPESNPKNSANLTTSLIKHAVKOTQOTLTGHQY-WKIAAMKLSSESKAKI 73
Db 44 FFFTLHLETPK--PSISSNLNPRETMEAVSLTCDPEPDASYLWMNGOSLPMTHSLKL 101
QY 74 SET 76
Db 102 SET 104

RESULT 40

US-08-619-812-2
Sequence 2, Application US/08619812
Patent No. 6100066
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: THEISEN, MICHAEL
APPLICANT: HARLAND, RICHARD J.
APPLICANT: RHOX, CLEMENT R.
TITLE OF INVENTION: VACCINES FOR HAEMOPHILUS SOMNUS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET

```

: CITY: PALO ALTO
: STATE: CALIFORNIA
: COUNTRY: UNITED STATES OF AMERICA
: ZIP: 94301
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/619,812
: FILING DATE: 15-MAR-1996
: CLASSIFICATION: 435
:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 08/038,719
: FILING DATE: 29-MAR-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: ROBINS, ROBERTA L.
: REGISTRATION NUMBER: 33,208
: REFERENCE/DOCKET NUMBER: 9000-0019.20
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 617-8999
: TELEFAX: (415) 327-3231
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 247 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: US-08-619-812-2

```

```

Query Match          9.28; Score 62.5; DB 3; Length 247;
Best Local Similarity 24.88; Pred. No. 11;
Matches 32; Conservative 14; Mismatches 60; Indels 23; Gaps 5;

QY 3 KLVYVATLMTAFILASCASPESNPKRSSANLITSLIKHAVKQTCQTQLTGHQYWKIAA 62
   | | : : | | : | | : | | : | | : | | : | | : | | : | | : | |
Db 5 KSLVGTIVASTVLIAAC-----NEKNKAETTPTEPTVAETQ-AQPDVVG-----KT 51

QY 63 MKLSSESKAKISSTACGCVADKAPAVSLTELTAAINP-----NARTEVAQKIYRHS LKP 118
   ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 52 ETTSES-----TAIENTOSDAQEKTETTSVETTSPTAAGNTOPESQEKVSEKSET 105

QY 119 CMLETVNAF 127
   : | : | |
Db 106 VVOEILNQF 114

```

Search completed: April 28, 2003, 16:05:10
 Job time : 20 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 16:04:32 ; Search time 19 Seconds
(without alignments)
565.126 Million cell updates/sec

Title: US-10-048-196-2

Perfect score: 676
Sequence: 1 NMKILYVATMTAFETLASC.....SLKPCMLETVNAFVPTTR 134

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications-AA:*

- 1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB pep:*
- 2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB pep:*
- 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB pep:*
- 4: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB pep:*
- 5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB pep:*
- 6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB pep:*
- 7: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB pep:*
- 8: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB pep:*
- 9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB pep:*
- 10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB pep:*
- 11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB pep:*
- 12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB pep:*
- 13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB pep:*
- 14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78.5	11.6	200	9 US-09-738-626-5845	Sequence 5845, App
2	76.5	11.3	386	10 US-09-864-761-35720	Sequence 35720, A
3	76.5	11.3	605	10 US-09-801-368-428	Sequence 428, App
4	72	10.7	2368	10 US-09-815-242-5635	Sequence 5635, App
5	72	10.7	2368	10 US-09-815-242-12389	Sequence 12389, A
6	71.5	10.6	722	9 US-10-029-495-5	Sequence 5, App11
7	71.5	10.6	737	9 US-10-029-806-731	Sequence 731, App
8	71.5	10.6	824	9 US-10-029-495-6	Sequence 6, App11
9	70	10.4	562	10 US-09-801-368-82	Sequence 82, App1
10	68.5	10.1	448	10 US-09-864-761-44230	Sequence 44230, A
11	68	10.1	2340	9 US-10-184-644-379	Sequence 379, App
12	68	10.1	2340	9 US-10-184-634-379	Sequence 379, App
13	67.5	10.0	1169	10 US-09-801-368-106	Sequence 106, App
14	66.5	9.8	878	9 US-10-254-534-2	Sequence 2, App11
15	66	9.8	206	10 US-09-815-242-10522	Sequence 10522, A
16	66	9.8	423	10 US-09-815-242-13560	Sequence 13560, A
17	66	9.8	1257	9 US-10-184-644-365	Sequence 365, App
18	66	9.8	1257	9 US-10-184-634-365	Sequence 365, App
19	66	9.8	2434	10 US-09-815-242-5835	Sequence 5835, App

20	66	9.8	6281	10 US-09-815-242-12996	Sequence 12996, A
21	65.5	9.7	1356	9 US-10-077-111-10	Sequence 10, App1
22	65.5	9.7	1571	9 US-10-184-644-609	Sequence 609, App
23	65.5	9.7	1571	9 US-10-123-155-97	Sequence 97, App1
24	65.5	9.7	1571	9 US-10-123-155-545	Sequence 545, App
25	65.5	9.7	1571	9 US-10-184-634-609	Sequence 609, App
26	65	9.6	2209	9 US-09-902-841-1903	Sequence 1903, App
27	65	9.6	2209	9 US-09-849-626-1903	Sequence 1903, App
28	65	9.6	2209	9 US-10-017-754-1903	Sequence 1903, App
29	64.5	9.5	2849	9 US-10-123-155-285	Sequence 285, App
30	64.5	9.5	3170	9 US-10-184-644-249	Sequence 249, App
31	64.5	9.5	3170	9 US-10-184-634-249	Sequence 249, App
32	64	9.5	270	9 US-09-738-626-3943	Sequence 3943, App
33	64	9.5	276	10 US-09-939-980-516	Sequence 516, App
34	64	9.5	513	9 US-10-149-819-9	Sequence 9, App11
35	64	9.5	1167	9 US-10-147-026-8	Sequence 8, App11
36	64	9.5	1325	10 US-09-864-761-35612	Sequence 35612, A
37	64	9.5	1342	9 US-10-172-620-16	Sequence 16, App1
38	64	9.5	3871	9 US-10-184-644-347	Sequence 347, App
39	64	9.5	3871	9 US-10-184-634-347	Sequence 347, App
40	64	9.5	5795	10 US-09-815-242-12610	Sequence 12610, A
41	63.5	9.4	152	9 US-09-769-952-12	Sequence 12, App1
42	63.5	9.4	276	9 US-09-769-952-2	Sequence 2, App11
43	63.5	9.4	285	9 US-09-769-952-14	Sequence 14, App1
44	63.5	9.4	1419	9 US-10-123-155-517	Sequence 517, App
45	63.5	9.4	1554	9 US-10-184-644-297	Sequence 297, App

ALIGNMENTS

RESULT 1
US-09-738-626-5845
Sequence 5845, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SANCOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOHI, AKIHIRO
APPLICANT: IKEDA, MASATO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738, 626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5845
LENGTH: 200
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5845

Query Match 11.6%; Score 78.5; DB 9; Length 200;
Best Local Similarity 23.5%; Pred. No. 0.54;
Matches 27; Conservative 28; Mismatches 47; Indels 13; Gaps 4;

OY 1 NMKILYVATMTAFETLASCSTPESNPKNSANLITSLIKHAYKOCQOQLGHOYKI 60
DB 3 LFKATVATVYAAALASACSSSDSSSTSSASDAATQVPTAEELN----- 56
OY 61 AAMLSSEKAKISE--TACGCVADKAPFAVSLTELTAAINPNAETVEAKIV 112

Db 57 AIAVATDEAPLEKVKTVQ--SENAPL--FETMQAKVSGAEFQVGSVL 107

RESULT 2

US-09-864-761-35720 ; Sequence 35720, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharon G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 FILE REFERENCE: Aecmca-x-1
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US/09/864,761
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: GB 24263,6
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 35720
 LENGTH: 386
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AP000511.1
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.7
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
 OTHER INFORMATION: SWISSPROT HIT: P28968, EVALUATE 2.00e-05
 US-09-864-761-35720

Query Match 11.3%; Score 76.5; DB 10; Length 386;
 Best Local Similarity 27.0%; Pred. No. 2.1;
 Matches 37; Conservative 18; Mismatches 59; Indels 23; Gaps 6;

QY 7 VTATLMTAFTLASCSTPESNPKNSSANLTTSLIKHAKVOTCOTOLGHOYMIAMKLS 66
 Db 125 VSATVPRKNTSPVITSTPSTAPMTASKMTTVA-SKTAATSTI-TSLPTVFTTSKITAG 182
 QY 67 SE-----SKAKISETAGCGVADKAPAVSLT--ELTTAANINARTEVAKIVRHS 115
 Db 183 SEIPTASTSDSATTAISTASTGTTVESAPSTAPPTPAETTTASVPTTSTGSENGHH- 241
 QY 116 LKPCMLETVNAFVPTT 132
 Db 242 -----TVSS--VPPT 249

RESULT 3

US-09-801-368-428
 ; Sequence 428, Application US/09801368
 Patent No. US20020128250A1
 GENERAL INFORMATION:
 APPLICANT: Busby, Robert
 APPLICANT: Call, Brian
 APPLICANT: Hecht, Peter
 APPLICANT: Holtzman, Doug
 APPLICANT: Madden, Kevin
 APPLICANT: Maxon, Mary
 APPLICANT: Milne, Todd
 APPLICANT: No. US20020128250A1man, Thea
 APPLICANT: Royer, John
 APPLICANT: Salama, Sofie
 APPLICANT: Sherman, Amir
 APPLICANT: Silva, Jeff
 APPLICANT: Summers, Eric
 TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
 FILE REFERENCE: 109272.147
 CURRENT FILING DATE: 2001-03-07
 PRIOR APPLICATION NUMBER: US 09/487,558
 PRIOR FILING DATE: 2000-01-19
 PRIOR APPLICATION NUMBER: US 60/160,587
 PRIOR FILING DATE: 1999-10-20
 NUMBER OF SEQ ID NOS: 440
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 428
 LENGTH: 605
 TYPE: PRT
 ORGANISM: Saccharomyces cerevisiae
 US-09-801-368-428

Query Match 11.3%; Score 76.5; DB 10; Length 605;
 Best Local Similarity 22.4%; Pred. No. 3.8;
 Matches 30; Conservative 24; Mismatches 71; Indels 9; Gaps 3;

QY 7 VTATLMTAFTLASCSTPESN-----PKNSSANLTTSLIKHAKVOTCOTOLGHOYMIAMKLS 61
 Db 154 LSTSTPTPLTTASTSTPTTPTDITSALEPTTSTKLSTSTPTSTSTSTSS--TST 211
 QY 62 AKKLSSEKAKISETAGCGVADKAPAVSLTETLT--AAINPNARTEVAKIVRHSLKRC 119
 Db 212 TVSVSTSTSTSTSTSTSTSTSTSTSSSSSTPTTSSAISTSTSTSTSTSTSS 271
 QY 120 MLETVNAFVPTT 133
 Db 272 SAPTSSNTPTPT 285

RESULT 4

US-09-815-242-5635
 ; Sequence 5635, Application US/09815242
 Patent No. US20020061569A1

```

: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Kari L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: TITLE OF INVENTION: Identification of Essential Genes In
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: NUMBER OF SEQ ID NOS: 1410
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5635
: LENGTH: 2368
: TYPE: PRT
: ORGANISM: Staphylococcus aureus
US-09-815-242-5635

Query Match          10 7%; Score 72; DB 10; Length 2368;
Best Local Similarity 29.7%; Pred. No. 71;
Matches 33; Conservative 11; Mismatches 33; Indels 34; Gaps 5;

QY      9  ATLMTAFILASCASPEESNPK-----NSSANLTTLIKHAVKQCGOTLGHQYKIAM 63
      ||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Db      2006  ATGETTATSAANSATDDANDKRPQANNSSADISTN-----SPYNDNDVT----- 2049

QY      64  KLSSESKAKISSETACGY-----ADKAPFAVSLT--ELTPAAINPNARTE 106
      ||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Db      2050  -----SKPEVESTNNGTGDKPATEADNMTPEASATINNNTTATNENAPTE 2055

RESULT 5
US-09-815-242-12389
: Sequence 12389, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Kari L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: TITLE OF INVENTION: Identification of Essential Genes In
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26

```

```

      PRIOR APPLICATION NUMBER: 60/242,578
      PRIOR FILING DATE: 2000-10-23
      PRIOR APPLICATION NUMBER: 60/253,625
      PRIOR FILING DATE: 2000-11-27
      PRIOR APPLICATION NUMBER: 60/257,931
      PRIOR FILING DATE: 2000-12-22
      PRIOR APPLICATION NUMBER: 60/269,308
      PRIOR FILING DATE: 2001-02-16
      NUMBER OF SEQ ID NOS: 14110
      SOFTWARE: FastSeq for Windows Version 4.0
      SEQ ID NO 12389
      LENGTH: 2368
      TYPE: PRT
      ORGANISM: Staphylococcus aureus
US-09-815-242-12389

Query Match                               10.7%; Score 72; DB 10; Length 2368;
Best Local Similarity 29.7%; Pred. No. 71;
Matches 33; Conservative 11; Mismatches 33; Indels 34; Gaps 5

QY      9 ATLMTAFLASCASPESNPK-----NSSANLITSLIKHAVKOTCOTLTGHQYKIAM 63
Db       2006 ATGETTATSNASSATDNDKPQANNSSADSTN-----SPTMDNDVT----- 2049
      |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      64 KLSSESKAKISTACGY-----ADKAPEAVSL--ELTTAINPAPARTE 106
Db       2050 -----SKPEVESTNGTTDKPATETADNPATPAESATNNNSTTATENMAPTE 2095
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 6
US-10-029-495-5
Sequence 5, Application US/100294.95
Publication No. US20030027255A1
GENERAL INFORMATION:
APPLICANT: Drmanac, Radoje T.
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crain, Birgit
APPLICANT: Dickson, Mark C.
APPLICANT: Jones, Lee William
APPLICANT: Ballinger, Dennis G.
APPLICANT: Xue, Aйдong
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi,Vinod
TITLE OF INVENTION: STEM CELL MAINTENANCE FACTOR MATERIALS AND METHODS
FILE REFERENCE: 28110/35905A
CURRENT APPLICATION NUMBER: US/10-029,495
CURRENT FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: 09/378,667
PRIOR FILING DATE: 1999-08-20
PRIOR APPLICATION NUMBER: 09/687,527
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 722
TYPE: PRT
ORGANISM: Caenorhabditis elegans
FEATURE:
OTHER INFORMATION: ciw11
US-10-029-495-5

Query Match                               10.6%; Score 71.5; DB 9; Length 722;
Best Local Similarity 21.1%; Pred. No. 17;
Matches 39; Conservative 20; Mismatches 49; Indels 73; Gaps 7;

QY      5 LYTATLTMTAETLASCAST-----PESPKNSSANLITSLIKHAVKOTCOTQ 51
Db       482 LYHSTLT-KGTVACAVSTTSNDFTQFSQTRPHENPTQLGNNL-----HFVRSKL-- 533
      ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      52 LTGHQYKIAMKISS-----ESKKISETACGCYADKAPEAVSLTE' 93
Db       534 ----QYDNDNDKTPLPSRLILYRDGAGDGQIPYIRNKTEVKALDCAVDATDKAAEISNRVO 589
      ::::~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~:

```



```

RESULT 10
US-09-864-761-44230
; Sequence 44230, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44230
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005666.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.42
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.5
; OTHER INFORMATION: SWISSPROT HIT: Q01484, EVALU6 7.00e-04
; OTHER INFORMATION: EST_HUMAN HIT: AA312125.1, EVALU6 1.00e-41
US-09-864-761-44230

```

```
Oy      24 PSBPNKSSANLTTSLIKHAVQOCOTQLTGHQYWKIA-----AMKISSESKAKISETAC 78
```

Db 14 PETGPTRESSRHSSPHLKQKQTQTATLASHSRQ-WKVSYPREAFARGSESK---SGREY 69
Qy 79 GCVADKAPRAVSLTELTATAINPNA RTEVAQ 109
70 GHHAASSIPDAKSTHQLSFDQKNNIQSDILE 100

```

RESULT 11
US-10-184-644-379
; Sequence 379, Application US/10184644
; Publication No. US20050044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P343061C27
; CURRENT APPLICATION NUMBER: US/10/184, 644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 379
; LENGTH: 2340
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-644-379

```

Query Match	10.1%;	Score 68;	DB 9;	Length 2340;
Best Local Similarity	23.9%;	Pred. No. 1.9e+02;		
Matches 26;	Conservative 10;	Mismatches 53;	Indels 20;	Gaps 2.

```

Oy      8 TATLMTAFTTASCASCPESNPKNSSANLTTSLIKHAVKOTQOTLTGHQWIKIAMKISS 67
        | | : | : : | | : | | |
Db    2140 TCTACATTTAGACCTATATAGAAATTTAGATATATGTATGTTCTGCAT-----TA 2187

```

QY	68	ESKAKISETAC-----GCVADKAPEAVSLTELTTAAINPNARTEVA	108
		:	
Db	2188	TAATATGCTACCACATGTAGCAATAATTACAATATTTTATTTAAATAAAA	2236

RESULT 12
US-10-184-634-379
; Sequence 379, Application US/10184634
; Publication No. US20030068684A1

```

? GENERAL INFORMATION:
? APPLICANT: Baker, Kevin P.
? APPLICANT: Chen, Jian
? APPLICANT: Desnoyers, Luc
? APPLICANT: Goddard, Audrey
? APPLICANT: Godowski, Paul J.
? APPLICANT: Gurney, Austin L.
? APPLICANT: Pan, James
? APPLICANT: Smith, Victoria
? APPLICANT: Watanabe, Colin K.
? APPLICANT: Wood, William I.
? APPLICANT: Zhang, Zemin
? TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
? FILE REFERENCE: P3430R1C217
? CURRENT APPLICATION NUMBER: US/10/184,634
? CURRENT FILING DATE: 2002-06-28
? Prior Application removed - See File Wrapper or Palm
? NUMBER OF SEQ ID NOS: 612
? SEQ ID NO 379

```

LENGTH: 2340
TYPE: DNA
ORGANISM: Homo Sapien
US-10-184-634-379

Query Match
Best Local Similarity 10.1%; Score 68; DB 9; Length 2340;
Matches 26; Conservative 10; Mismatches 53; Indels 20; Gaps 2;

OY 8 TATLTAFTLASCASPESNPKSSANLTTSLIKHAKQTCOTOLTGHOYKIAAMKLSS 67
DB 2140 TCGAGATTGACTATATACAAATTGATGATGTCGCAT-----TA 2187
OY 68 ESKAKISERAC-----GCVADKAPKAVSLTETLTAIPNAKTEVA 108
DB 2188 TATATGTCACACATGTACCAATTAATTACAAATTTTATTAATAATAA 2236

RESULT 13

US-09-801-368-106
Sequence 106, Application US/09801368
Patent No. US20020128250A1
GENERAL INFORMATION:

APPLICANT: Busby, Robert
APPLICANT: Call, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin
APPLICANT: Maxon, Mary
APPLICANT: Moline, Todd
APPLICANT: No. US20020128250A1man, Thea
APPLICANT: Royer, John
APPLICANT: Salama, Sofie
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
CURRENT APPLICATION NUMBER: US/09/801.368
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: US 60/160,587
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 440
SOFTWARE: PatentIn version 3.0
SEQ ID NO 106
LENGTH: 1169
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-801-368-106

Query Match
Best Local Similarity 25.0%; Score 67.5; DB 10; Length 1169;
Matches 29; Conservative 20; Mismatches 38; Indels 29; Gaps 6;

OY 13 TATLTAFTLASCASPESNPKSSANLTTSLIKHAKQTCOTOLTGHOYKIAAMKLSS 71
DB 938 TLITVSSCS-----NSCSNTVSSAVSTATTITTINGITTEYTWCPLSATELTIVSKL 990
OY 72 K-----ISEAC--GCVADKAPKAVSLTETLTAIPNAKTEVA 108
DB 991 ESEKTLITVSSCSGVCSEFASPAIVST--ATATVADVIVSTVSWSPQATNKLA 1044

RESULT 14

US-10-254-534-2
Sequence 2, Application US/10254534
Publication No. US20030046730A1
GENERAL INFORMATION:
APPLICANT: EK, Bo
APPLICANT: KHOSNOODI, Jamshid
APPLICANT: LARSSON, Clas-Tomas

APPLICANT: LARSSON, Hakan
APPLICANT: RASK, Lars

TITLE OF INVENTION: STARCH BRANCHING ENZYME II OF POTATO
FILE REFERENCE: 003300-486
CURRENT APPLICATION NUMBER: US/10/254,534
CURRENT FILING DATE: 2002-09-26
PRIOR APPLICATION NUMBER: US/09/087,277
PRIOR FILING DATE: 1998-05-29
PRIOR APPLICATION NUMBER: PCT/SE96/01558
PRIOR FILING DATE: 1996-11-28
PRIOR APPLICATION NUMBER: SE 9504272-7
PRIOR FILING DATE: 1995-11-29
PRIOR APPLICATION NUMBER: SE 9601506-0
PRIOR FILING DATE: 1996-04-19
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2

LENGTH: 878
TYPE: PRT
ORGANISM: Unknown
FEATURE:

OTHER INFORMATION: Description of Unknown Organism:beli gene (branching enzyme II
US-10-254-534-2

Query Match
Best Local Similarity 9.8%; Score 66.5; DB 9; Length 878;
Matches 28; Conservative 21; Mismatches 43; Indels 23; Gaps 4;

OY 26 SNPKSSANLTTSLIKHAKQTCOTOLTGHOYKIAAMKLSSERAKKISERAC----- 78
DB 24 SNGDRNANXSVFLKKHSLSR-----KILAEKSYNSSESRSTVAASGVLP 71
OY 79 GCVADKAPKAVSLTETLTAIPNAKTEVAQKIVRHSLSKPCMLTETVNAFYPTT 132
DB 72 GTQSDSSSSSTDQFEFTERSPENSASPSTDVDSITVEMARQ---IKTENDVEPSS 123

RESULT 15

US-09-815-242-10522
Sequence 10522, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011a
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10522
LENGTH: 206

```
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10522

Query Match
Best Local Similarity 24.7%; Pred. No. 13;
Matches 23; Conservative 16; Mismatches 42; Indels 12; Gaps 2;

OY 24 PESNPKSSANL--TSLIKHAVKOTCOTOLTGHOYWKIAMKLSSEKAKISETAGCV 81
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 65 PDKAPKDTANILATKELVHFVNEDELQMN-----QTSAPLAALISIDIFGI 114
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |

OY 82 ADKAPKAVSLTELTTAINDNARTENAOKIVRH 114
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 115 ETIASQTVAVPAIKAPIRMEARLHOYPIANH 147

RESULT 16
US-09-815-242-13560
; Sequence 13560, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13560
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13560

Query Match
Best Local Similarity 9.8%; Score 66; DB 10; Length 423;
Matches 22; Conservative 11; Mismatches 27; Indels 10; Gaps 2;

OY 65 LSSESRKISSETACG---CVADKAP-----EAVSLTEITTAINDNARTENAOKIVRH 114
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 231 LSKESVYKVERFAGGAEVTAQVPHFSEKTEALLTQSNKMNPLLESBRAYIE 290
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |

OY 115 SLKPCMLETV 124
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 291 GLKSGVITVI 300

RESULT 17
US-10-184-644-365
; Sequence 365, Application US/10184644
```

```
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 365
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-634-365

Query Match
Best Local Similarity 9.8%; Score 66; DB 9; Length 1257;
Matches 27; Conservative 9; Mismatches 43; Indels 24; Gaps 3;

OY 8 TATIMTAFTLASCASTPESNPKSSANLTTSLIKHAVKOTCOTOLTGHOYWKIAMKLSS 67
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 897 TTTTATTTATGCGCTTGGAATGTTCACTTA---AATGACATTTT-----AAATAG 945
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |

OY 68 ESKAKISETAC-----GCVADKAPKAVSLTELTTA 97
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 946 TTTATGTATACATCTGATGAAAGCAAGCTAATATGTGTTA 988
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |

RESULT 18
US-10-184-634-365
; Sequence 365, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 365
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-634-365

Query Match
Best Local Similarity 9.8%; Score 66; DB 9; Length 1257;
Matches 27; Conservative 9; Mismatches 43; Indels 24; Gaps 3;

OY 8 TATIMTAFTLASCASTPESNPKSSANLTTSLIKHAVKOTCOTOLTGHOYWKIAMKLSS 67
```

Db 897 TTTTATTATGCTTGGAATGCTTCACTTA-----AATGACATTTT-----AAATPAG 945
OY 68 ESKAKISETAC-----GCVADKADPAVSLTELTAA 97
Db 946 TTATGTATACATCTGATGAAAGCAAGCAAGCTAATATGTCTTA 988

RESULT 19

US-09-815-242-5835
Sequence 5835, Application US/09815242
Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA 011A

CURRENT APPLICATION NUMBER: US/09/815,242

PRIOR APPLICATION NUMBER: 2001-03-21

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5835

LENGTH: 2434

TYPE: PRT

ORGANISM: Staphylococcus aureus

US-09-815-242-5835

Query Match
Best Local Similarity 9.8%; Score 66; DB 10; Length 2434;
Matches 27; Conservative 17; Mismatches 43; Indels 20; Gaps 5;

OY 17 LASCSTPESNPKNSANLTTSLIKHAVKOTCOTQLTGHOYKIAAKLSSSEKA-----71
Db 1175 IISATNNPEMNP--DTINOKASQVNSA-----KSALNGDE--KLAARKOTAKSDIGRLTD 1225

OY 72 --KISFTACGCVADKADPAVSLTELTAAINPNARTEVAQKIVRHSI 116
Db 1226 LNNMORTAANAEDVQAPNLAAY-----TAARKKATSLMTAMGNLKHAI 1268

RESULT 20

US-09-815-242-12996

Sequence 12996, Application US/09815242
Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA 011A

CURRENT APPLICATION NUMBER: US/09/815,242

PRIOR APPLICATION NUMBER: 2001-03-21

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 12996

LENGTH: 6281

TYPE: PRT

ORGANISM: Staphylococcus aureus

US-09-815-242-12996

Query Match
Best Local Similarity 9.8%; Score 66; DB 10; Length 6281;
Matches 27; Conservative 17; Mismatches 43; Indels 20; Gaps 5;

OY 17 LASCSTPESNPKNSANLTTSLIKHAVKOTCOTQLTGHOYKIAAKLSSSEKA-----71
Db 1954 IISATNNPEMNP--DTINOKASQVNSA-----KSALNGDE--KLAARKOTAKSDIGRLTD 2004

OY 72 --KISFTACGCVADKADPAVSLTELTAAINPNARTEVAQKIVRHSI 116
Db 2005 LNNMORTAANAEDVQAPNLAAY-----TAARKKATSLMTAMGNLKHAI 2047

RESULT 21

US-10-077-111-10

Sequence 10, Application US/10077111
Publication No. US20020187492A1

GENERAL INFORMATION:

APPLICANT: Todderud, C. Gordon

APPLICANT: Finger, Joshua N.

APPLICANT: Rillema, J111

FILE REFERENCE: 3053-4114US2

CURRENT APPLICATION NUMBER: US/10/077,111

PRIOR APPLICATION NUMBER: 2002-02-15

PRIOR FILING DATE: 2001-05-29

PRIOR APPLICATION NUMBER: 60/294,181

PRIOR FILING DATE: 2001-05-29

PRIOR APPLICATION NUMBER: 60/269,366

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 25

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 10

LENGTH: 1356

TYPE: PRT

ORGANISM: Podospora anserina

FEATURE:

OTHER INFORMATION: beta transducin-like protein encoded by the

US-10-077-111-10

Query Match
Best Local Similarity 9.7%; Score 65.5; DB 9; Length 1356;
Matches 31; Conservative 16; Mismatches 55; Indels 25; Gaps 4;

OY 19 SCSTPESNPKNSANLTTSLIKHAVKOTCOTQLTGHOY-----KIAAKLSS 67


```
Db 1182 SVASPDQORVASGSDKIRIMPTASGCTQLEHGHWQSVAFSPDQORVASG--SS 1239
OY 68 ESKAKISTACGCVADKAPAEVSLT---ELTTAIPNARTEVAOKIVRHSLSKCMLET 123
Db 1240 DNTKIMPTASGCTQTLNVGSTATCLSPDTNMIYNTN-----IGRIQIATATMES 1291
OY 124 VNAETVP 130
Db 1292 LNOQLSP 1298

RESULT 22
US-10-184-644-609
; Sequence 609, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C27
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 609
; LENGTH: 1571
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-644-609

Query Match 9.7%; Score 65.5; DB 9; Length 1571;
Best Local Similarity 26.0%; Pred. No. 2.1e+02;
Matches 33; Conservative 9; Mismatches 62; Indels 23; Gaps 4;

OY 9 ATLMTAFT--LASCASTPESNPKSSANLITSLKHAKOTCOTQOLGHOYWKIAMKLS 66
Db 1073 ATCTTAATTGAGACCACTACTGTATTAAGATTGTTGTAATATCTTCTG-----CTATTGG 1127
OY 67 SESKAKISTACGCVADKAPAEVSLTETLTPAIPNARTEVAOKIVRHSLSKCMLETVNA 126
Db 1128 ATATATTATTATTAGTTAATA-----TATTATTATTATTTTTG-----CTATTAA 1171
OY 127 FIVPTTT 133
Db 1172 TGTATTT 1178

RESULT 23
US-10-123-155-97
; Sequence 97, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
```

```
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 97
; LENGTH: 1571
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-97

Query Match 9.7%; Score 65.5; DB 9; Length 1571;
Best Local Similarity 26.0%; Pred. No. 2.1e+02;
Matches 33; Conservative 9; Mismatches 62; Indels 23; Gaps 4;

OY 9 ATLMTAFT--LASCASTPESNPKSSANLITSLKHAKOTCOTQOLGHOYWKIAMKLS 66
Db 1073 ATCTTAATTGAGACCACTACTGTATTAAGATTGTTGTAATATCTTCTG-----CTATTGG 1127
OY 67 SESKAKISTACGCVADKAPAEVSLTETLTPAIPNARTEVAOKIVRHSLSKCMLETVNA 126
Db 1128 ATATATTATTATTAGTTAATA-----TATTATTATTATTTTTG-----CTATTAA 1171
OY 127 FIVPTTT 133
Db 1172 TGTATTT 1178

RESULT 24
US-10-123-155-545
; Sequence 545, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 545
; LENGTH: 1571
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-545

Query Match 9.7%; Score 65.5; DB 9; Length 1571;
```

Best Local Similarity 26.0%; Pred. No. 2.1e+02;
Matches 33; Conservative 9; Mismatches 62; Indels 23; Gaps 4;

QY 9 ATLMTAFT--LASCSTPESNPKNSSANLTTSLIKHAVKOTQOTOLGHQYWKIAAMKLS 66
DB 1073 ACCTTAATGAGACCACTTGTATAGATTTTGTAAATATCTTCTG-----CTATTGG 1127
QY 67 SESKAKISFTACGCVAADKAPENAVSLTELTTAAINPNAFTEVAOKIVRHSLSKPCMLETVNA 126
DB 1128 AATATATTATTTAGTTAATA-----TATTATTTATTTTGTG-----CTATTAA 1171
QY 127 FIVPTTT 133
DB 1172 TGTATTT 1178

RESULT 25
US-10-184-634-609
; Sequence 609, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P343081C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 609
; LENGTH: 1571
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-184-634-609

Query Match 9.7%; Score 65.5; DB 9; Length 1571;
Best Local Similarity 26.0%; Pred. No. 2.1e+02;
Matches 33; Conservative 9; Mismatches 62; Indels 23; Gaps 4;

QY 9 ATLMTAFT--LASCSTPESNPKNSSANLTTSLIKHAVKOTQOTOLGHQYWKIAAMKLS 66
DB 1073 ACCTTAATGAGACCACTTGTATAGATTTTGTAAATATCTTCTG-----CTATTGG 1127
QY 67 SESKAKISFTACGCVAADKAPENAVSLTELTTAAINPNAFTEVAOKIVRHSLSKPCMLETVNA 126
DB 1128 AATATATTATTTAGTTAATA-----TATTATTTATTTTGTG-----CTATTAA 1171
QY 127 FIVPTTT 133
DB 1172 TGTATTT 1178

RESULT 26
US-09-902-941-1903
; Sequence 1903, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Mainerakis, Margarita

APPLICANT: Carter, Derrick
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: McNeabb, Andria
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.478C17
CURRENT APPLICATION NUMBER: US/09/902,941
CURRENT FILING DATE: 2001-07-10
NUMBER OF SEQ ID NOS: 2002
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1903
LENGTH: 2209
TYPE: PRT
ORGANISM: Homo sapiens
US-09-902-941-1903

Query Match 9.6%; Score 65; DB 9; Length 2209;
Best Local Similarity 21.0%; Pred. No. 3.8e+02;
Matches 30; Conservative 20; Mismatches 41; Indels 52; Gaps 6;

QY 30 NSSANLTTSLIKH-----AVKOTCOTOL-----TGHQYW 58
DB 1546 NININQALSLIKHLSYRISPPVDELYQYMLEHVITLPSAOTRPLIFFGTAQNEW 1605
QY 59 KIAAMKLSSES-----KAKIS-----ETACGCVAADK-APEAVSLTELTTAAINPN 102
DB 1606 KILSTELSESFPYLLLSIKMKFSLDLYSTAKHVFEEKLPKLLKLTQAKSSTLTK 1665
QY 103 ARTEVAOKIVRHSLSKPCMLETVN 125
DB 1666 EITKITQ-----TIESCLISYV 1683

RESULT 27
US-09-849-626-1903
; Sequence 1903, Application US/09849626
; Publication No. US20020197659A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aljun
; APPLICANT: Wang, Tonglong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1903
; LENGTH: 2209
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-626-1903

Query Match 9.6%; Score 65; DB 9; Length 2209;
Best Local Similarity 21.0%; Pred. No. 3.8e+02;
Matches 30; Conservative 20; Mismatches 41; Indels 52; Gaps 6;

QY 30 NSSANLTTSLIKH-----AVKOTCOTOL-----TGHQYW 58
DB 1546 NININQALSLIKHLSYRISPPVDELYQYMLEHVITLPSAOTRPLIFFGTAQNEW 1605
QY 59 KIAAMKLSSES-----KAKIS-----ETACGCVAADK-APEAVSLTELTTAAINPN 102
DB 1606 KILSTELSESFPYLLLSIKMKFSLDLYSTAKHVFEEKLPKLLKLTQAKSSTLTK 1665
QY 103 ARTEVAOKIVRHSLSKPCMLETVN 125


```
Sequence 249, Application US/10184634
: Publication No. US20030068684A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jian
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P3430R1C217
: CURRENT APPLICATION NUMBER: US/10/184,634
: PRIOR FILING DATE: 2002-06-28
: Prior Application removed - See file wrapper or Palm
: NUMBER OF SEQ ID NOS: 612
: SEQ ID NO 249
: LENGTH: 3170
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-184-634-249

Query Match          9.5%; Score 64.5; DB 9; Length 3170;
Best Local Similarity 27.6%; Pred. No. 6.9e+02;
Matches 27; Conservative 10; Mismatches 50; Indels 11; Gaps 3;

QY 8 TATMTATATLASCSTPESNPKNSANLTTSLIKHAVKOTCOTOLTGHOYKIAMKLSS 67
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2895 TTCTATATTAATACATATTA-----TTATATTACA-ATTGGTTCTG-----CAATATT 2943

QY 68 ESKAKISFTACGCVAADKAPAVSLTELTAAINPART 105
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2944 TTCTATGTCCACCCTTTTAAATAATTATATTGAGCT 2981

RESULT 32
US-09-738-626-3943
: Sequence 3943, Application US/09738626
: Publication No. US20020197605A1
: GENERAL INFORMATION:
: APPLICANT: NAKAGAMA, SATOSHI
: APPLICANT: MIZOGUCHI, HIROSHI
: APPLICANT: ANDO, SEIKO
: APPLICANT: HAYASHI, MIKIRO
: APPLICANT: OCHIAI, KEIKO
: APPLICANT: YOKOI, HARUHIKO
: APPLICANT: TATEISHI, NAKO
: APPLICANT: SENOH, AKIHIRO
: APPLICANT: IKEDA, MASATO
: APPLICANT: OZAKI, AKIO
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-125
: CURRENT APPLICATION NUMBER: US/09/738,626
: PRIOR FILING DATE: 2000-12-18
: PRIOR APPLICATION NUMBER: JP 99/377484
: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: JP 00/159162
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: JP 00/280988
: NUMBER OF SEQ ID NOS: 7059
: SOFTWARE: PatentIn ver. 3.0
: SEQ ID NO 3943
: LENGTH: 270
: TYPE: PRT
: ORGANISM: Corynebacterium glutamicum
US-09-738-626-3943
```

```
Query Match          9.5%; Score 64; DB 9; Length 270;
Best Local Similarity 23.3%; Pred. No. 31;
Matches 27; Conservative 10; Mismatches 45; Indels 34; Gaps 3;

QY 19 SCASPTES-----NPKNSANLTTSLIKHAVKOTCOTOL---TGHOYKIAA 62
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 99 SISCTPESTIQLOGLDIGBDNPNVADHLTTLTATDTEFIATIQQMAEATGG----- 151

QY 63 MKLSSEKAKISFTACGCVAADKAPAVSLTELTAAINPARTVAVOKIYRHSKLP 118
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 152 -----NSTSGFAEQLIQELIKVTDIISDANQTIIEVFPTDGAARASLHP 196

RESULT 33
US-09-939-980-516
: Sequence 516, Application US/09939980
: Patent No. US20020082234A1
: GENERAL INFORMATION:
: APPLICANT: Black, Michael
: Burnham, Martin
: Hodgson, John
: Knowles, David
: Lonetto, Michael
: Pratt, Julie
: Reichard, Richard
: Rosenberg, Martin
: Ward, Judith
: TITLE OF INVENTION: No. US20020082234A1el Prokaryotic Polynucleotides,
: NUMBER OF SEQUENCES: 534
: CORRESPONDENCE ADDRESS:
: ADDRESSSEE: SmithKline Beecham Corporation
: STREET: 709 Swedeland Road
: CITY: King of Prussia
: STATE: PA
: COUNTRY: USA
: ZIP: 19406-0939
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/939,980
: FILING DATE: 27-Aug-2001
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/936,165
: FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Gimmil, Edward R
: REGISTRATION NUMBER: 38,891
: REFERENCE/DOCKET NUMBER: P50549
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-270-4478
: TELEFAX: 610-270-5090
: TELEX: <Unknown>
: INFORMATION FOR SEQ ID NO: 516:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 276 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: Protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 516:
US-09-939-980-516

Query Match          9.5%; Score 64; DB 10; Length 276;
Best Local Similarity 25.2%; Pred. No. 32;
Matches 27; Conservative 16; Mismatches 44; Indels 20; Gaps 4;

QY 17 LASCSTPESNPKNSANLTTSLIKHAVKOTCOTOLTGHOYKIAMKLSSSKA----- 71
```

```
Db 106 IISAXNPEMP-----DIXOKASOVNSAKSALXGDE--KLAAMQOTAKSDIGRVTD 156
OY 72 --KISFACGCVADKAPKAVSLFELTAATPNMRTFVAKIVHSL 116
Db 157 LNNMORTAXNNEVDAPXLAIV---TAAKNKATSLMTAMGVKXHAL 199

RESULT 34
US-10-149-819-9
; Sequence 9, Application US/10149819
; Publication No. US20030044913A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AN-YOUNG, Janice
; APPLICANT: BUREFORD, Neil
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: PF-0760 PCT
; CURRENT APPLICATION NUMBER: US/10/149,819
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
; PRIOR FILING DATE: 1999-12-10; 1999-12-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030044913A1 2215706CD1
US-10-149-819-9

Query Match
Best local Similarity 9.5%; Score 64; DB 9; Length 513;
Matches 33; Conservative 15; Mismatches 44; Indels 38; Gaps 5;

OY 25 ESNPKXSSANLT-----TSLIK-----HAYKQTQOTLTGHQYKTIAMKLSSSKA 71
Db 334 ESSPNHSLNHEVADDSHLEKANLIELEDDSHSGKRGIPHSLSGIODPIIARMSICSEDK 393
OY 72 KISFACGCVADKAPKAVSLFELTAATPNMRTFVAKIVHSL 116
Db 394 SPSE--CSLIASSPEENWPAKCAKAYINLRIPSTVTLNNSAPANRANQNFDEMEGIRETS 451
OY 109 QKIVRHSKLP 118
Db 452 QVILRPSSP 461

RESULT 35
US-10-147-026-8
; Sequence 8, Application US/10147026
; Publication No. US20030003538A1
; GENERAL INFORMATION:
; APPLICANT: Dietrich, Paul Shartzter
; APPLICANT: Huang, Chiao-Chain
; APPLICANT: Johnson, Carl D.
; APPLICANT: Sangameswaran, lakshmi
; TITLE OF INVENTION: Neuropathic Pain Genes: Compositions
; FILE REFERENCE: ROCH-006
; CURRENT APPLICATION NUMBER: US/10/147,026
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 60/155,702
```

```
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/189,931
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1167
; TYPE: PRT
; ORGANISM: rattus sp.
US-10-147-026-8

Query Match
Best local Similarity 9.5%; Score 64; DB 9; Length 1167;
Matches 29; Conservative 23; Mismatches 49; Indels 34; Gaps 4;

OY 13 TAFILASCSTPESNPKNSSANLTSLIKHAKOTQOTLTGHQYKTIAMKLSSSKAK 72
Db 673 TGISGLIAGSSIONPKVSLPLPARLLQGLGRSSQIHAASSQ-----THVSSQAQAA 727
OY 73 ISEFACGCVADKAPKAVSLFELTAATPNMRTFVAKIVHSL 116
Db 728 ASSHALG--TSEADASSLTQVTYVHSHAVQONVYSPLOATISKSTQNPVYKLSNPNOL 785
OY 106 EVAOKIVRHSKLP 120
Db 786 SCSSQLKTSKPLM 800

RESULT 36
US-09-864-761-35612
; Sequence 35612, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Neomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
```

FILE REFERENCE: P3430R1C217

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 16:02:12 : Search time 17 Seconds
(without alignments)
757.765 Million cell updates/sec

Title: US-10-048-196-2

Perfect score: 676

Sequence: 1 MKKILYVTATLMTAFTLASC.....SLKPCMLETVNAEIVPTTTR 134

Scoring table: BL0SUM62

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

PIR_73: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	118	17.5	123	2 F81834	probable lipoprote
2	116	17.2	123	2 G81127	hypothetical prote
3	88.5	13.1	142	2 A95132	hypothetical prote
4	82	12.1	179	2 C68853	hypothetical prote
5	80	11.8	262	2 A49220	immunoreactive lip
6	79.5	11.7	4377	2 A55575	ankyrin 3, long sp
7	79	11.4	354	2 T46740	microfilament shea
8	77	11.3	967	2 S66852	hypothetical prote
9	76.5	11.3	414	2 A29835	Sali protein - Esc
10	76.5	11.3	605	2 S48940	hypothetical prote
11	76	11.2	535	2 T05863	hypothetical prote
12	76	11.2	995	2 S50358	hypothetical prote
13	75.5	11.2	1041	2 S55862	probable membrane
14	75.5	11.2	189	2 AE0560	conserved hypotet
15	75	11.1	1203	2 T17415	mycelial surface a
16	74.5	11.0	3507	2 T34513	hypothetical prote
17	74	10.9	298	1 MGNZRL	major surface glyco
18	73.5	10.9	1758	2 S57015	probable purine nu
19	72.5	10.7	160	2 E81844	hypothetical prote
20	72.5	10.7	424	2 H96963	dehydroxylase (im
21	72.5	10.7	928	2 S64350	hypothetical prote
22	72.5	10.7	1104	2 S59310	probable membrane
23	72.5	10.7	1122	2 G64887	probable tail fibe
24	72.5	10.7	2718	2 A23475	G surface protein
25	71.5	10.6	722	2 T30395	hypothetical prote
26	71.5	10.6	824	2 T20351	hypothetical prote
27	71	10.5	171	2 AH1643	hypothetical prote
28	71	10.5	743	2 D84545	probable salt-indu
29	71	10.5	900	2 C64232	alanine-tRNA ligas

30	70.5	10.4	676	2 S41217	hypothetical prote
31	70.5	10.4	696	2 S39827	SKTs protein - yea
32	70.5	10.4	1161	2 S57180	probable membrane
33	70.5	10.4	1229	2 T25697	hypothetical prote
34	70	10.4	171	2 A11280	hypothetical prote
35	70	10.4	402	2 E86185	hypothetical prote
36	70	10.4	562	2 B41035	chitinase (EC 3.2.
37	70	10.4	562	2 S50371	chitinase (EC 3.2.
38	70	10.4	1260	2 S60896	agglutinin-like pr
39	69.5	10.3	190	2 E71508	probable methylase
40	69.5	10.3	235	2 PC2022	mucin like protein
41	69.5	10.3	495	2 A40985	probable zinc-prot
42	69.5	10.3	2479	2 F87386	conserved hypotet
43	69	10.2	189	2 G81423	cytolethal distend
44	69	10.2	910	2 S73361	dead homolog prote
45	69	10.2	993	2 G84632	hypothetical prote

ALIGNMENTS

```

RESULT 1
F81834
Probable lipoprotein NMA1439 [imported] - Neisseria meningitidis (strain Z2491 serogr
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: F81834
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: F81834
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-123 <PAR>
A:Cross-references: GB:A162756; GB:A157959; NID:g7380091; PIDN:CA884675.1; PID:g738
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA1439
C:Superfamily: Neisseria meningitidis hypothetical protein NMB1047

Query Match 17.5%; Score 118; DB 2; Length 123;
Best Local Similarity 23.1%; Pred. No. 0.00023;
Matches 28; Conservative 29; Mismatches 58; Indels 6; Gaps 3;

OY 1 MKKILYVTATLMTAFTLASCSTPES--NPKSSANLITSLIKHAVKQTCQQLGHQYW 58
DB 1 MKKILSI---LPVALILGGCAAGGNTGSLDGTGKGSIYKMAVESQCRALNKRSEW 57
OY 59 KIAAMKLSSESKAKISRTACGCVADKAPDAVSLTELTTAIPNPA RTEVAQIVRHSILKP 118
DB 58 RLTALAMSAEKQAEWENKICACVQAEAFNQLTGNDV--HQMIDPSTRNQAALATRTKYSA 116
OY 119 C 119
DB 117 C 117

RESULT 2
G81127
Hypothetical protein NMB1047 [imported] - Neisseria meningitidis (strain MC58 serogro
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: G81127
R:Telletlin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Maignani, V.; Piazza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Mokron, E.R.; Rappuoli, R.;
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20157551; PMID:10710307
A:Accession: G81127

```

ankyrin 5, long splice form - human
N:Alternate names: ankyrin G
C:Species: Homo sapiens (man)
C:Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 20-Sep-1999
C:Accession: A55575
R:Kordelli, E.; Lambert, S.; Bennett, V.
J. Biol. Chem. 270, 2352-2359, 1995
A>Title: Ankyrin-G. A new ankyrin gene with neural-specific isoforms localized at the
A:Reference number: A55575; MUID:95138209; PMID:7836469
A:Accession: A55575
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-4377 <KOR>
A:Cross-references: GB:U13616; NID:g608024; PIDD:AAA64834.1; PID:g608025
C:Genetics:
A:Gene: GDB:ANK3
A:Cross-references: GDB:424503; OMIM:600465

A:Map position: 10q21-10q21
C:Superfamily: unassigned ankryrin repeat proteins; ankryrin repeat homology; EGF homology
C:Keywords: alternative splicing; peripheral membrane protein
F:7-105/Domain: ankryrin repeat homology <AN01>
F:106-138/Domain: ankryrin repeat homology <AN02>
F:139-171/Domain: ankryrin repeat homology <AN03>
F:172-200/Domain: ankryrin repeat homology <AN04>
F:201-233/Domain: ankryrin repeat homology <AN05>
F:234-266/Domain: ankryrin repeat homology <AN06>
F:267-299/Domain: ankryrin repeat homology <AN07>
F:300-332/Domain: ankryrin repeat homology <AN08>
F:333-365/Domain: ankryrin repeat homology <AN09>
F:366-398/Domain: ankryrin repeat homology <AN10>
F:399-431/Domain: ankryrin repeat homology <AN11>
F:432-464/Domain: ankryrin repeat homology <AN12>
F:465-497/Domain: ankryrin repeat homology <AN13>
F:498-530/Domain: ankryrin repeat homology <AN14>
F:531-563/Domain: ankryrin repeat homology <AN15>
F:564-596/Domain: ankryrin repeat homology <AN16>
F:597-629/Domain: ankryrin repeat homology <AN17>
F:630-662/Domain: ankryrin repeat homology <AN18>
F:663-695/Domain: ankryrin repeat homology <AN19>
F:696-728/Domain: ankryrin repeat homology <AN20>
F:729-761/Domain: ankryrin repeat homology <AN21>
F:762-794/Domain: ankryrin repeat homology <AN22>
F:795-827/Domain: ankryrin repeat homology <AN23>

Query Match 11.8%; Score 79.5; DB 2; Length 4377;
Best Local Similarity 26.7%; Pred. No. 67;
Matches 28; Conservative 18; Mismatches 52; Indels 7; Gaps 4;

QY 14 AFTLASCASTPESNP-KNSANLTLTLIKHAVKOTQOTLGHQYWKIAMKLSSEKAK 72
Db 3900 ALTTSSCVDRKSHLPKNTFRD---NI--AVRKACATQOG-QPEGKAKQLPSKLPVK 3953
QY 73 ISETAGCVAADKAPKPEAVSLTELTTAIPNAPKEVNAOKIVRHSIK 117
Db 3954 VRSTCVTTTTTTTATTTTTTTTTTCTVAKRSQLEKVKHSIE 3998

RESULT 7
T46740
microfilament sheath protein SHP3 [Imported] - Litomosoides sigmodontis
C:Species: Litomosoides sigmodontis
C>Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000
C:Accession: T46740
R:Hiemann, J.
Submitted to the EMBL Data Library, October 1999
A:Reference number: Z23140
A:Accession: T46740
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-354 <HIR>
A:Cross-references: EMBL:U54556; PIDN:AAB82017.2
C:Genetics:
A:Gene: shp3
A:Introns: 19/3

Query Match 11.7%; Score 79; DB 2; Length 354;
Best Local Similarity 24.1%; Pred. No. 4.5;
Matches 32; Conservative 20; Mismatches 75; Indels 6; Gaps 4;

QY 2 MKILYVATLMTAFTLASCASTPESNPNKNSANLTLTLIKHAVKOTQOTLGHQYWKIA 61
Db 1 MTEVLDVIVAFELL--CV--TPRASASSSTSTTPAK--TSTTTTAKTTSKTTKTT 55
QY 62 AKLISESEKAKISETACGCVADKAPKPEAVSLTELTTAIPNAPKEVNAOKIVRHSIKPCML 121
Db 56 TVKSTTTTST 114
QY 122 EYVNAFIVPTTR 134
Db 115 STTSTTKPTTSK 127

RESULT 8

hypothetical protein YOL155c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein 00419; protein ADF1001
C:Species: Saccharomyces cerevisiae
C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002
C:Accession: S66852; S66854; S67325; S70380
R:Arino, J.; Casamayor, A.; Gamo, F.J.; Gancedo, C.; Lafuente, M.J.; Aldea, M.; Casas
submitted to the Protein Sequence Database, July 1996

A:Reference number: S66852
A:Accession: S66852
A:Molecule type: DNA
A:Residues: 1-967 <GAI>
A:Cross-references: EMBL:Z74897; NID:91420063; PID:e251930; PID:91420064; MIPS:YOL155
A:Experimental source: strain S288C
R:Gallion, L.; Dujon, B.
Submitted to the Protein Sequence Database, July 1996

A:Reference number: S66854
A:Accession: S66854
A:Molecule type: DNA
A:Residues: 1-967 <GAI>
A:Cross-references: EMBL:Z74897; NID:91420063; PID:e251930; PID:91420064; MIPS:YOL155
A:Experimental source: strain S288C
R:Gamo, F.J.; Lafuente, M.J.; Casamayor, A.; Aldea, M.; Casas, C.; Arino, J.; Herrero,
submitted to the EMBL Data Library, July 1995

A:Description: Analysis of the DNA sequence of a 15500 bp fragment of the left arm of
pen reading frames.
A:Reference number: S67324
A:Accession: S67325
A:Molecule type: DNA
A:Residues: 1-164, 'STSTSGSSATSSSVSGTSATSSGSSASGSS', 166-186, 'V', 188-967 <GAM>
A:Cross-references: EMBL:X89715; NID:91177620; PID:e190152; PID:91177622
R:Gamo, F.J.; Lafuente, M.J.; Casamayor, A.; Arino, J.; Aldea, M.; Casas, C.; Herrero
Yeast 12, 709-714, 1996

A:Title: Analysis of the DNA sequence of a 15,500 bp fragment near the left telomere
and two new open reading frames.
A:Reference number: S70379; MUID:96405919; PMID:8810044
A:Accession: S70380
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 55-164, 'STSTSGSSATSSSVSGTSATSSGSSASGSS', 166-186, 'V', 188-286, 745-773;
A:Cross-references: EMBL:X89715
C:Genetics:
A:Cross-references: SGD:S0005515
A:Map position: 15L
A:Note: YOL155c

Query Match 11.4%; Score 77; DB 2; Length 967;
Best Local Similarity 24.3%; Pred. No. 22;
Matches 34; Conservative 21; Mismatches 57; Indels 28; Gaps 7;

QY 7 VTAATL--MTAFTLASCAST-----PESNKNSSANLTLTLIKHAVKOTQOT-LT 53
Db 709 ITTTVPCCSTTATITTSODETGHVTTSTGVAFETVSSKSTTYTVTHCONGNTKTVT 768
QY 54 GHQYWKIAAMKLSSEKAKISETAC---GC-----VADKAPKPEAVSLTELTTAIPNAPKE 106
Db 769 SECEPEESATTTSPKSTTYTVTHCDNCGNCTKTVTSEARDA-----TTTVSP--KTY 820
QY 107 VQAKIVRHSIKPCMLETVNA 126
Db 821 TTAFTVTOCDNDGCSKTVTS 840

RESULT 9

AZ29835
SalI protein - Escherichia coli plasmid PED208 (fragment)
C:Species: Escherichia coli
C>Date: 04-Aug-1988 #sequence_revision 04-Aug-1988 #text_change 23-Mar-1993
C:Accession: A29835
R:Finlay, B.B.; Paranchych, W.

J. Bacteriol. 166, 713-721, 1986
 A:Title: Nucleotide sequence of the surface exclusion genes *trac* and *tracT* from the IncF-
 A:Reference number: A29835; MUID:86223783; PMID:3011738
 A:Accession: A29835
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1414 <FIN>
 C:Genetics:
 A:Genome: plasmid

Query Match 11.3%; Score 76.5; DB 2; Length 414;
 Best Local Similarity 29.4%; Pred. No. 9.4;
 Matches 32; Conservative 16; Mismatches 46; Indels 15; Gaps 5;

QY 26 SNPKNSANL-----TSLIKHAKVQTCOTQLTGHQWKIA--AMKLSSEKAKISERA 77
 Db 192 SHTDNASASLANQSLSTSSLSKQASQNDATRSHESQLASVYENNSASINONYSOEF 251
 QY 78 CGVADKAP-EAVSLTELTAAINPNARTE--VAQKIVRHSLKPCMLE 122
 Db 252 VGVTYSKRPGEA---DQLSDAASPEVRAERDRLAQAFAVEDMKRQLQ 297

RESULT 10
 S48940
 hypothetical protein YH028w - yeast (Saccharomyces cerevisiae)
 C:Species: Saccharomyces cerevisiae
 C:Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 29-Oct-1999
 C:Accession: S48940
 R:Favella, T.
 submitted to the EMBL Data Library, June 1994
 A:Description: The sequence of S. cerevisiae cosmid 9196.
 A:Reference number: S46794
 A:Accession: S48940
 A:Molecule type: DNA
 A:Residues: 1-605 <FAV>
 A:Cross-references: EMBL:U11583; NID:92289854; PIDN:AAB65040.1; PID:92289857; MIPS:YH028
 C:Genetics:
 A:Gene: SCD:MSC4
 A:Cross-references: SCD:S0001020; MIPS:YH028w
 A:Map position: 8L
 C:Keywords: transmembrane protein

Query Match 11.3%; Score 76.5; DB 2; Length 605;
 Best Local Similarity 22.4%; Pred. No. 14;
 Matches 30; Conservative 24; Mismatches 71; Indels 9; Gaps 3;

QY 7 VTATLMTAFILASCASPESN-----PKNSANLTTSLIKHAKVQTCOTQLTGHQWKIA 61
 Db 154 LRTSTTPTLTASTSTPTSDITSALEPTTSTKLSISIPSTSTSTSTSS--TST 211
 QY 62 AMKLSESKAKISERACGVADKAPAVSLTELT--AAINPNARTEVAQKIVRHSLKPC 119
 Db 212 TVSVTSTSTSTSTSTSTSTSTSSSSSTPTTSSAPISITSTSTSTSTSTSTSS 271
 QY 120 MEETVAATFVPTTT 133
 Db 272 SAPTSSSNTPTST 285

RESULT 11
 T05863
 hypothetical protein T29A15.70 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 13-Aug-1999
 C:Accession: T05863
 R:Bevan, M.; Van der Schueren, J.; Chung, Y.-J.; Voet, M.; Robben, J.; Volckaert, G.; Ho
 submitted to the Protein Sequence Database, March 1999
 A:Reference number: 215455
 A:Accession: T05863
 A:Molecule type: DNA
 A:Residues: 1-515 <BEV>
 A:Cross-references: EMBL:AL035602

A:Experimental source: cultivar Columbia; BAC clone T29A15
 C:Genetics:
 A:Map position: 4
 A:introns: 52/2; 133/3; 183/3; 203/1; 234/1; 267/1; 297/3; 317/3; 394/1; 451/3
 A:Note: T29A15.70

Query Match 11.2%; Score 76; DB 2; Length 515;
 Best Local Similarity 29.3%; Pred. No. 13;
 Matches 27; Conservative 15; Mismatches 44; Indels 6; Gaps 3;

QY 16 TLASCASP-ESNPKNSANLTTSLT--KHAKVQTCOTQLTGHQWKIAAMKLSSEKAK 72
 Db 411 TNGACSKRKESDIYEGSVSTENAVESKNAATEIDATLTQEKRESIETRKKEETKED 470
 QY 73 ISETACGVADKAPAVSLTELTAAINPNAR 104
 Db 471 SEEAR---KAEPTEPAVKAEKTSSETEPPAQ 499

RESULT 12
 S50358
 hypothetical protein YIL169c - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein Y19402.07c
 C:Species: Saccharomyces cerevisiae
 C:Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 19-Apr-2002
 C:Accession: S50358
 R:Lye, G.; Bowman, S.; Churcher, C.
 submitted to the EMBL Data Library, December 1994
 A:Reference number: S50349
 A:Accession: S50358
 A:Molecule type: DNA
 A:Residues: 1-995 <LYE>
 A:Cross-references: GB:247047; EMBL:Z46921; NID:9603997; PID:9604006; GSPDB:GN00009;
 C:Genetics:
 A:Gene: MIPS:YIL169c
 A:Cross-references: SCD:S0001431
 A:Map position: 9L

Query Match 11.2%; Score 76; DB 2; Length 995;
 Best Local Similarity 20.6%; Pred. No. 28;
 Matches 35; Conservative 24; Mismatches 59; Indels 52; Gaps 6;

QY 7 VTATLMTAFILASCASPES-----NPKNSANLTTSLIKHAKVQTCOTQLT 53
 Db 686 ITTVPSSSTTATTTSCDEGCHVSTGAVVTEVTSKSYTTATVTHCDNCGNTKTVT 745
 QY 54 GHQYWKIAAMKLSSEKAKISERAC---GC---VADKAPAVSLT----- 92
 Db 746 SECKERTSATTAASKITVTVVHCDNCGNTKTVTSEAPATTTTVSSQSYTTATVTH 805
 QY 93 -----ELTTAAINPNARTEVAQKIVRHSLKPCMLETVNA 126
 Db 806 CDDNGCKTKTIVTSEAPATTTTVSP--KYTTATVTHQCDNCGSTKTVTS 853

RESULT 13
 S55862
 probable membrane protein YNL327w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein N0320
 C:Species: Saccharomyces cerevisiae
 C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jul-2000
 C:Accession: S55862; S63308; S51288
 R:Maitah, M.; Nicaud, J.M.; Levesque, H.; Galliardin, C.
 yeast 11, 567-572, 1995
 A:Title: Sequencing analysis of a 15.4 kb fragment of yeast chromosome XIV identifies
 A:Reference number: S55859; MUID:95373280; PMID:7645347
 A:Accession: S55862
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-1041 <MAF>
 A:Cross-references: EMBL:Z46259; NID:9633655; PIDN:CAA6371.1; PID:9633659
 R:Maitah, M.; Nicaud, J.M.; Levesque, H.; Galliardin, C.
 submitted to the Protein Sequence Database, April 1996

A:Reference number: S63287
 A:Accession: S63308
 A:Molecule type: DNA
 A:Residues: 1-1041 <MAN>
 A:Cross-references: EMBL:771603; NID:q1302444; PID:e239572; PID:q1302445; MIPS:YNL327W
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGD:EGT2
 A:Cross-references: SGD:S0005271; MIPS:YNL327W
 A:Map position: 14L
 C:Keywords: transmembrane protein
 F:4-20/Domain: transmembrane #status predicted <TM1>
 F:1025-1041/Domain: transmembrane #status predicted <TM2>

Query Match 11.2%; Score 76; DB 2; Length 1041;
 Best Local Similarity 25.0%; Pred. No. 30;
 Matches 30; Conservative 22; Mismatches 56; Indels 12; Gaps 3;
 Oy 24 PESNPKNSANLTTSLIKHAVKOTCOTQLTGHQYWK-----IAAMKLSSSEKAKISETA 77
 Db 912 PVSYIKGSKTSLST---EESVAGYSTVDSDAQAEHTNLVAIDTLKTSFQKATATEVC 968
 Oy 78 CGGVADKAPAVSLTETLTAINPNARTEVAOKIVRHS---LKPCMLETVNAFIPTTTR 134
 Db 969 VTCALSSPSHSATLDAGTTLSTPTSSSTLSTTTWTSSSTIKRPSISTYSGAAGQLTIR 1028

RESULT 14
 AE0560
 conserved hypothetical 11protein SRY0509 [imported] - Salmonella enterica subsp. enter
 C:Species: Salmonella enterica subsp. enterica serovar typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
 C:Accession: AE0560
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,
 Th, T.; Conington, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Garra, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A:Reference number: AB0502; PMID:11677608
 A:Accession: AE0560
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-189 <PAR>
 A:Cross-references: GB:AL51382; PIDN:CAD08925.1; PID:q16501737; GSPDB:GN00176
 C:Genetics:
 A:Gene: SRY0509

Query Match 11.2%; Score 75.5; DB 2; Length 189;
 Best Local Similarity 26.7%; Pred. No. 4.9;
 Matches 39; Conservative 21; Mismatches 47; Indels 39; Gaps 9;
 Oy 2 MKLIYVTTMTAFTLASCA-----STPESNPKNSANLTTSLIKHAVKOTCOTQLTGH 55
 Db 1 MKLVHIVGLAVIALSLACAKSADIQPPAPNP-NKSTIRANSTIQ-----QPVSC- 51
 Oy 56 QYV---KIA-----AMKLSESEKAK-----ISETACGCVADKAPAVSLTETLTAIN 100
 Db 52 TWVIMOKAALPDVAVLTITSLDASLADAPSKYLSOKAVTEKQAPFSLV-LPFPNSDIO 110
 Oy 101 PNARTEVAOKI-----VRHSLEK 118
 Db 111 PNARILLSAATVDNKLVEITDSVKP 136
 RESULT 15
 T17415
 mycelial surface antigen CSa1 precursor - yeast (Candida albicans)
 C:Species: Candida albicans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T17415
 R:Bourdonnais, Y.; Lamarre, C.; Deslauriers, N.

submitted to the EMBL Data Library, July 1998
 A:Description: Expression cloning of the Candida albicans CSa1 gene encoding a myceli
 A:Reference number: Z16775
 A:Accession: T17415
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1203 <BOU>
 A:Cross-references: EMBL:AF080221; NID:q3406797; PID:g3406798; PIDN:AAC29486.1
 A:Experimental source: strain ATCC32354
 C:Genetics:
 A:Gene: CSa1
 F:18-1203/Product: mycelial surface antigen CSa1 #status predicted <MAT>
 F:1-17/Domain: signal sequence #status predicted <SIG>

Query Match 11.1%; Score 75; DB 2; Length 1203;
 Best Local Similarity 24.5%; Pred. No. 43;
 Matches 36; Conservative 25; Mismatches 54; Indels 32; Gaps 6;
 Oy 8 TATLMTAFTLASCASTPESNPKNSANLTTSLIKHAVK-----QTCOTQLT 53
 Db 371 TSSETSSQQLSITSAIDSSSTSTSTIRASINGFADKLYDLPKCAKPCMFQNT 430
 Oy 54 G---HOYWKIAMKLSSSEKAKISETACGVAD--KAPAVSLTETLTAIN---PNART 105
 Db 431 GTPCPYVDACCLCVMQFAGAIQ---SCVADSCGQDIVSVSLGTSVCSVAGVAPY 486
 Oy 106 EVAOKIVRHSLEKPCMLETVNAFIPTT 132
 Db 487 WMLPASVRSLSL-----SVAAATAVPTS 507

RESULT 16
 T34513
 hypothetical protein ZK783.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T34513
 R:Favell, A.; Vaudin, M.
 submitted to the EMBL Data Library, August 1994
 A:Description: The sequence of C. elegans cosmid ZK783.
 A:Reference number: Z21536
 A:Accession: T34513
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-3507 <FAV>
 A:Cross-references: EMBL:U13646; PIDN:AAC24418.1; GSPDB:GN00021; CESP:ZK783.1
 A:Experimental source: strain Bristol N2; clone ZK783
 C:Genetics:
 A:Gene: CESP:ZK783.1
 A:Map position: 3
 A:Introns: 14/1; 48/2; 84/1; 196/3; 303/1; 381/1; 586/1; 605/1; 1175/3; 1207/1; 1409/
 3504/1

Query Match 11.0%; Score 74.5; DB 2; Length 3507;
 Best Local Similarity 21.6%; Pred. No. 1.6e+02;
 Matches 30; Conservative 30; Mismatches 50; Indels 29; Gaps 4;
 Oy 13 TAFTLASCASTPESNPKNSANLTTSLIKHAVKOTCOTQLTGHQWKIAMKLSS----- 67
 Db 2449 TSNVLSSESTPSSSKSVSSTEGI-----SVATSTFEKVPSESTISSVLEED 2498
 Oy 68 -----ESKAKISEKACGVADKAPAVSLTETLTAIN---PNARTEVAOKIVRHS 114
 Db 2499 LKTPPSPILETTTASSTSEPLETSLTVSVRIHELTTSENKVPKESSTTSSSSKRP 2558
 Oy 115 SLKPCMLETVNAFIPTT 133
 Db 2559 SQEPAGILT-STVVPPTSS 2576

RESULT 17
 MG2ZRL
 major surface glycoprotein G - human respiratory syncytial virus (strain Long)

C:Species: human respiratory syncytial virus
 C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
 C:Accession: A32703; S12279
 R:Johnson, P.R.; Spriggs, M.K.; Orlsted, R.A.; Collins, P.L.
 Proc. Natl. Acad. Sci. U.S.A. 84, 5625-5629, 1987
 A:Title: The G glycoprotein of human respiratory syncytial viruses of subgroups A and B
 A:Reference number: A32703; MUID:87289657; PMID:2441388
 A:Accession: A32703
 A:Molecule type: mRNA
 A:Residues: 1-298 <JON>
 A:Cross-references: GB:M17212; NID:9333940; PIDN:AAA47411.1; PID:9333941
 R:Garcia-Barrero, B.; Portella, A.; Delgado, T.; Lopez, J.A.; Melero, J.A.
 EMBO J. 9, 4181-4187, 1990
 A:Title: Frame shift mutations as a novel mechanism for the generation of neutralization
 A:Reference number: S12279; MUID:91065351; PMID:2249671
 A:Accession: S12279
 A:Molecule type: mRNA
 A:Residues: 1-298 <GAR>
 C:Superfamily: respiratory syncytial virus major surface glycoprotein G
 C:Keywords: glycoprotein; transmembrane protein
 F:41.63/Domain: transmembrane status predicted <TM>
 F:85.103,135,179,237,250,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status

Query Match 10.98; Score 74; DB 1; Length 298;
 Best Local Similarity 24.38; Pred. NO. 11;
 Matches 27; Conservative 20; Mismatches 50; Indels 14; Gaps 3;

OY 1 MKILVYATLMTAFLLASCASTESNPNKSSANLTSIKHA--VKQOTQOLTGQY 57
 Db 45 ILAVISTSLITLITLIFAS-----NKKVLTITLIDATSOIKNTPTLTLDPO 96
 OY 58 WKIAMALS---SESKAKISSETACGCVAADKAPKAVSLTETLTAIINPNT 105
 Db 97 LGISFSNLSSETTSQTTLASTTPGVKSNLOPTVKKTKNTTQTQPSKPT 147

RESULT 18
 S57015
 probable purine nucleotide-binding protein YJL225C - Yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein J0202; hypothetical protein L9122.1-a; hypotheti-
 cal protein Y19402.01c; protein HMD1054; protein HMD732
 C:Species: Saccharomyces cerevisiae
 C:Date: 08-Jul-1995 #sequence_revision 08-Sep-1995 #text_change 19-Apr-2002
 C:Accession: S57015; S51848; S50702; S45147; S50349; S31214; S59002; S48919; S69476; S58
 R:Vandenbol, M.; Durand, P.; Portetelle, D.; Hliger, F.
 Submitted to the Protein Sequence Database, September 1995
 A:Reference number: S56835
 A:Accession: S57015
 A:Molecule type: DNA
 A:Residues: 1-1758 <VAN>
 A:Cross-references: EMBL:249500; NID:91015618; PID:91017444; MIPS:YJL225C
 R:Vandenbol, M.; Durand, P.; Portetelle, D.; Hliger, F.
 Submitted to the EMBL Data Library, January 1995
 A:Description: Sequence analysis of yeast chromosome X left telomere.
 A:Reference number: S51848
 A:Accession: S51848
 A:Molecule type: DNA
 A:Residues: 1-1758 <VAN>
 A:Cross-references: EMBL:248148; NID:9663331; PID:9663332
 R:Vandenbol, M.; Durand, P.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hliger, F.
 yeast 10, 1657-1662, 1994
 A:Title: Sequence analysis of a 40.2 kb DNA fragment located near the left telomere of y
 A:Reference number: S50701; MUID:95242842; PMID:7725802
 A:Accession: S50701
 A:Molecule type: DNA
 A:Residues: 1-1758 <VAN>
 A:Cross-references: EMBL:2343098
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994
 A:Note: neither the complete nucleic acid sequence nor the complete translation are show
 R:Vandenbol, M.; Durand, P.; Dion, C.; Bolle, P.; Portetelle, D.; Hliger, F.
 submitted to the EMBL Data Library, June 1994

A:Description: Sequence analysis of a 40.1 kb DNA fragment located near the left telo
 A:Reference number: S45146
 A:Accession: S45146
 A:Molecule type: DNA
 A:Residues: 1-73, 'I', '75-84', 'Q', '86-348', 'T', '350-387', 'VC', '944-946', 'FNR', '950', 'Y', '952', 'RCF
 SYSLPVLSARY', '955', 'SMIOQ', '961', 'A', '963', 'NCT', '967-968', 'T', '970', 'RT', '973', 'HRS', '978', 'LRR
 A:Cross-references: EMBL:025841; EMBL:273537; NID:91370583; PID:91370584
 A:Note: the nucleotide sequence was translated assuming an intron after the first bas
 A:Accession: S58831
 A:Molecule type: DNA
 A:Residues: 1-73, 'I', '75-84', 'Q', '86-150', 'PCPG', '439', 'LS', '442', 'NSECO', '448', 'EROL', '453-455,
 R:Medler, H.; Medler, E.; Scharfe, M.; Wambutt, R.
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S64792
 A:Accession: S65339
 A:Molecule type: DNA
 A:Residues: 1-73, 'I', '75-84', 'Q', '86-150', 'PCPG', '439', 'LS', '442', 'NSECO', '448', 'EROL', '453-455,
 A:Cross-references: EMBL:273537; MIPS:YPR202w
 A:Experimental source: strain S286C (AB972)
 R:Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sas
 submitted to the EMBL Data Library, May 1995
 A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces
 A:Reference number: S56186
 A:Accession: S56191
 A:Molecule type: DNA
 A:Residues: 1-65, 'N', '67-73', 'I', '75-84', 'Q', '86-150', 'PCPG', '439', 'LS', '442', 'NSECO', '448', 'EROL
 A:Cross-references: EMBL:050617; NID:9836685; PIDN:BAAO9177.1; PID:01009815; PID:9836
 R:Murakami, Y.

submitted to the EMBL Data Library, December 1994

A:Reference number: S62230

A:Accession: S62273

A:Molecule type: DNA

A:Residues: 1-65, 'N', 67-73, 'I', 75-84, 'Q', 86-150, 'PCPG', 439, 'LS', 442, 'NSECO', 448, 'EROL', 449, 'R', 451-452, 'R', 454-455, 'R', 457-458, 'R', 460-461, 'R', 463-464, 'R', 466-467, 'R', 469-470, 'R', 472-473, 'R', 475-476, 'R', 478-479, 'R', 481-482, 'R', 484-485, 'R', 487-488, 'R', 490-491, 'R', 493-494, 'R', 496-497, 'R', 499-500, 'R', 502-503, 'R', 505-506, 'R', 508-509, 'R', 511-512, 'R', 514-515, 'R', 517-518, 'R', 520-521, 'R', 523-524, 'R', 526-527, 'R', 529-530, 'R', 532-533, 'R', 535-536, 'R', 538-539, 'R', 541-542, 'R', 544-545, 'R', 547-548, 'R', 550-551, 'R', 553-554, 'R', 556-557, 'R', 559-560, 'R', 562-563, 'R', 565-566, 'R', 568-569, 'R', 571-572, 'R', 574-575, 'R', 577-578, 'R', 580-581, 'R', 583-584, 'R', 586-587, 'R', 589-590, 'R', 592-593, 'R', 595-596, 'R', 598-599, 'R', 601-602, 'R', 604-605, 'R', 606-607, 'R', 609-610, 'R', 612-613, 'R', 615-616, 'R', 618-619, 'R', 621-622, 'R', 624-625, 'R', 627-628, 'R', 630-631, 'R', 633-634, 'R', 636-637, 'R', 639-640, 'R', 642-643, 'R', 645-646, 'R', 648-649, 'R', 651-652, 'R', 654-655, 'R', 657-658, 'R', 660-661, 'R', 663-664, 'R', 666-667, 'R', 669-670, 'R', 672-673, 'R', 675-676, 'R', 678-679, 'R', 681-682, 'R', 684-685, 'R', 687-688, 'R', 690-691, 'R', 693-694, 'R', 696-697, 'R', 699-700, 'R', 702-703, 'R', 705-706, 'R', 708-709, 'R', 711-712, 'R', 714-715, 'R', 717-718, 'R', 720-721, 'R', 723-724, 'R', 726-727, 'R', 729-730, 'R', 732-733, 'R', 735-736, 'R', 738-739, 'R', 741-742, 'R', 744-745, 'R', 747-748, 'R', 750-751, 'R', 753-754, 'R', 756-757, 'R', 759-760, 'R', 762-763, 'R', 765-766, 'R', 768-769, 'R', 771-772, 'R', 774-775, 'R', 777-778, 'R', 780-781, 'R', 783-784, 'R', 786-787, 'R', 789-790, 'R', 792-793, 'R', 795-796, 'R', 798-799, 'R', 801-802, 'R', 804-805, 'R', 806-807, 'R', 809-810, 'R', 812-813, 'R', 815-816, 'R', 818-819, 'R', 821-822, 'R', 824-825, 'R', 827-828, 'R', 830-831, 'R', 833-834, 'R', 836-837, 'R', 839-840, 'R', 842-843, 'R', 845-846, 'R', 848-849, 'R', 851-852, 'R', 854-855, 'R', 857-858, 'R', 860-861, 'R', 863-864, 'R', 866-867, 'R', 869-870, 'R', 872-873, 'R', 875-876, 'R', 878-879, 'R', 881-882, 'R', 884-885, 'R', 887-888, 'R', 890-891, 'R', 893-894, 'R', 896-897, 'R', 899-900, 'R', 902-903, 'R', 904-905, 'R', 906-907, 'R', 909-910, 'R', 912-913, 'R', 915-916, 'R', 918-919, 'R', 920-921, 'R', 922-923, 'R', 924-925, 'R', 926-927, 'R', 928-929, 'R', 930-931, 'R', 932-933, 'R', 934-935, 'R', 936-937, 'R', 938-939, 'R', 940-941, 'R', 942-943, 'R', 944-945, 'R', 946-947, 'R', 948-949, 'R', 950-951, 'R', 952-953, 'R', 954-955, 'R', 956-957, 'R', 958-959, 'R', 960-961, 'R', 962-963, 'R', 964-965, 'R', 966-967, 'R', 968-969, 'R', 970-971, 'R', 972-973, 'R', 974-975, 'R', 976-977, 'R', 978-979, 'R', 980-981, 'R', 982-983, 'R', 984-985, 'R', 986-987, 'R', 988-989, 'R', 990-991, 'R', 992-993, 'R', 994-995, 'R', 996-997, 'R', 998-999, 'R', 1000-1001, 'R', 1002-1003, 'R', 1004-1005, 'R', 1006-1007, 'R', 1008-1009, 'R', 1010-1011, 'R', 1012-1013, 'R', 1014-1015, 'R', 1016-1017, 'R', 1018-1019, 'R', 1020-1021, 'R', 1022-1023, 'R', 1024-1025, 'R', 1026-1027, 'R', 1028-1029, 'R', 1030-1031, 'R', 1032-1033, 'R', 1034-1035, 'R', 1036-1037, 'R', 1038-1039, 'R', 1040-1041, 'R', 1042-1043, 'R', 1044-1045, 'R', 1046-1047, 'R', 1048-1049, 'R', 1050-1051, 'R', 1052-1053, 'R', 1054-1055, 'R', 1056-1057, 'R', 1058-1059, 'R', 1060-1061, 'R', 1062-1063, 'R', 1064-1065, 'R', 1066-1067, 'R', 1068-1069, 'R', 1070-1071, 'R', 1072-1073, 'R', 1074-1075, 'R', 1076-1077, 'R', 1078-1079, 'R', 1080-1081, 'R', 1082-1083, 'R', 1084-1085, 'R', 1086-1087, 'R', 1088-1089, 'R', 1090-1091, 'R', 1092-1093, 'R', 1094-1095, 'R', 1096-1097, 'R', 1098-1099, 'R', 1100-1101, 'R', 1102-1103, 'R', 1104-1105, 'R', 1106-1107, 'R', 1108-1109, 'R', 1110-1111, 'R', 1112-1113, 'R', 1114-1115, 'R', 1116-1117, 'R', 1118-1119, 'R', 1120-1121, 'R', 1122-1123, 'R', 1124-1125, 'R', 1126-1127, 'R', 1128-1129, 'R', 1130-1131, 'R', 1132-1133, 'R', 1134-1135, 'R', 1136-1137, 'R', 1138-1139, 'R', 1140-1141, 'R', 1142-1143, 'R', 1144-1145, 'R', 1146-1147, 'R', 1148-1149, 'R', 1150-1151, 'R', 1152-1153, 'R', 1154-1155, 'R', 1156-1157, 'R', 1158-1159, 'R', 1160-1161, 'R', 1162-1163, 'R', 1164-1165, 'R', 1166-1167, 'R', 1168-1169, 'R', 1170-1171, 'R', 1172-1173, 'R', 1174-1175, 'R', 1176-1177, 'R', 1178-1179, 'R', 1180-1181, 'R', 1182-1183, 'R', 1184-1185, 'R', 1186-1187, 'R', 1188-1189, 'R', 1190-1191, 'R', 1192-1193, 'R', 1194-1195, 'R', 1196-1197, 'R', 1198-1199, 'R', 1200-1201, 'R', 1202-1203, 'R', 1204-1205, 'R', 1206-1207, 'R', 1208-1209, 'R', 1210-1211, 'R', 1212-1213, 'R', 1214-1215, 'R', 1216-1217, 'R', 1218-1219, 'R', 1220-1221, 'R', 1222-1223, 'R', 1224-1225

```

C:Superfamily: Neisseria meningitidis hypothetical protein NMA1529

Query Match          10.7%; Score 72.5; DB 2; Length 160;
Best Local Similarity 28.4%; Pred. No. 8;
Matches 42; Conservative 18; Mismatches 55; Indels 33; Gaps 8;

OY 1 MKKIIYVTA-TLMTAFETLASCST-----PESNPRNSSANLTSLIKHAVQCOTQLTG 54
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2 MKKIIAVSLCLMT--AARADTYGYLAWONPONMADV-----QVVTTEEDSTKS 52
OY 55 HOYKIAAMKLSSEKAKISE---TAC-----GCVAADKAPAVSLTELTTTAINPNA 103
    : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
DB 53 EAFAELEAFCKQDPTLAGIAEDEPTGCRSVSYSLNNTCVALLPRAKMGKRVENAVLTSP 112

OY 104 R-TEVAQKIVRHSLK-----PCMLETV 124
    ||| ||| : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 113 RFTSVYQVALNCQICKKYGAGCGGLETV 140

RESULT 20
H96963
dihydroorotase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: H96963
R:Noiling, J.; Brenon, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
J. Dairy, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: H96963
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-424 <KUD>
A:Cross-references: GB:AE001437; PIDN:AAK78499.1; PID:q15023383; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC624
C:Genetics:
A:Gene: CAC0519
C:Superfamily: Bacillus dihydroorotase; Bacillus dihydroorotase homology

Query Match          10.7%; Score 72.5; DB 2; Length 424;
Best Local Similarity 23.0%; Pred. No. 24;
Matches 32; Conservative 24; Mismatches 46; Indels 37; Gaps 4;

OY 1 MKKIIYVTA-TLMTAFETL-SCASPESNPKNSSANLTSLIKHAVKQCTQLTGHOYWK 59
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 201 MGALEAEATVLTADALAVSSGAKINIQHSSKISLGIT----- 241
OY 60 IAAMKLSSEKAKISERACGVAADKAPAVSLTEL-----TTAINEPNAETVAOKIVR 113
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 242 ----KLAEMGMANI-----IAEATPQHFTSETEILNCGTNKAVNPPLAREDDRAIIV 290
OY 114 HSLKPCMLETVNAFLVPTT 132
    : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
DB 291 AALKDDTQIVATDHAAPT 309

RESULT 21
S64350
hypothetical protein YGR056w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein G4343
C:Species: Saccharomyces cerevisiae
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 20-Sep-1999
C:Accession: S64350
R:Entian, K.D.; Rose, M.; Koetter, P.; Roelmer, A.; Sehrsam, I.; Hempel, S.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64335
A:Accession: S64350
A:Molecule type: DNA
A:Residues: 1-928 <ENT>
A:Cross-references: EMBL:Z72841; NID:q1333068; PIDN:CAA97057.1; PID:e243414; PID:q1332
A:Experimental source: strain S288C
C:Genetics:

```


Species: *Caenorhabditis elegans*

Db 164 PNL-TCNTLLI 174

RESULT 29

C64232

alanine-RNA ligase (EC 6.1.1.7) - Mycoplasma genitalium
N:Alternate names: alanyl-tRNA synthetase

C:Species: Mycoplasma genitalium

C:Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 03-Jun-2002

C:Accession: C64232

R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;
M.; Fuhman, J.; Nguyen, D.; Uterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.
C.A.; Venter, J.C.
Science 270, 397-403, 1995

A:Title: The minimal gene complement of Mycoplasma genitalium.

A:Reference number: A64200; MUID:96026346; PMID:756993

A:Accession: C64232

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-900 <TIGR>

A:Cross-references: GB:U39710; GB:L43967; NID:g1045989; PID:g1045990; TIGR:MG292

A:Experimental source: strain G-37

C:Genetics:

A:Genetic code: SGC3

C:Superfamily: alanine-tRNA ligase

C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 10.5%; Score 71; DB 2; Length 900;

Best Local Similarity 24.1%; Pred. No. 77;

Matches 26; Conservative 23; Mismatches 49; Indels 10; Gaps 4;

QY 22 STPSNPNKNSANLTSLIKHAVKOTCOTLTGHQYKIAMKLSSESKAKISFTACGCV 81

Db 235 SVLQNSPTNFTDIFLKLK-IEAFCEPKYDPNSYFFPDQKVEQSYFRI-----I 286

QY 82 ADKPAEVSLELTAAINPNARFEVAKIYRHSLSKPCMLETVN-AFT 128

Db 287 ADHE-KAITFTISEGVLPGERMYVVRLLRLALACKKLQNLMAFI 333

RESULT 30

S41217

hypothetical protein YBL009v - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YBL0317

C:Species: Saccharomyces cerevisiae

C:Date: 28-Jan-1994 #sequence_revision 09-Sep-1994 #text_change 19-Apr-2002

C:Accession: S41217; S45742

R:Delaveau, T.; Jacq, C.; Perea, J.

Yeast 8, 761-768, 1992

A:Title: Sequence of a 12.7 kb segment of yeast chromosome II identifies a PDR-like gene

A:Reference number: S25326; MUID:93070613; PMID:1441753

A:Accession: S41217

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-676

A:Cross-references: EMBL:S47695; NID:g259049; PIDN:AAB23988.1; PID:g1680405

R:Delaveau, T.; Jacq, C.; Perea, J.

Submitted to the Protein Sequence Database, August 1994

A:Reference number: S45736

A:Accession: S45742

A:Molecule type: DNA

A:Residues: 1-676

A:Cross-references: EMBL:Z35770; NID:g535993; PIDN:CAA84828.1; PID:g535994; MIPS:YBL009w

A:Experimental source: strain S288C

C:Genetics:

A:Cross-references: SGD:S0000105

A:Map position: 2L

Query Match 10.4%; Score 70.5; DB 2; Length 676;

Best Local Similarity 22.6%; Pred. No. 62;

Matches 33; Conservative 18; Mismatches 48; Indels 47; Gaps 5;

QY 13 TATPLASCATPPSNPKNSANLT-----TSLIKHAVKOTCQ 49

Db 274 SGFIMPDHOSTKELNHHSSNSLSFRSLKHTKTSLSLNLKLVKRGKGTQGLNPIKKTQC 333

QY 50 TOLTGHQYW---KIAMKLSSESKAKISFTACGCVADKPAEVSLELTAAINPNARFE 106

Db 334 ISLPVDPQVSKDKIQLKLNKSTSLASLS-----EVPINLTLDVNDSTLQO 379

QY 107 VAG---KIYRHSLSKPCM---LETVN 125

Db 380 ILQLCVKTYLHDLREAGSLGFTLN 405

RESULT 31

S39827

SKT5 protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YBL0506; hypothetical protein YBL0519; hypoth

C:Species: Saccharomyces cerevisiae

C:Date: 16-Dec-1993 #sequence_revision 09-Sep-1994 #text_change 21-Jul-2000

C:Accession: S39827; S45796; S45483; PNO572; S37328; S42687

R:Schierens, B.; el Bakoury, M.; Vierendeels, F.; Dubois, E.; Messenguy, F.

Yeast 9, 1355-1371, 1993

A:Title: Sequencing and functional analysis of a 32 560 bp segment on the left arm of

A:Reference number: S39824; MUID:94205266; PMID:8154187

A:Accession: S39827

A:Molecule type: DNA

A:Residues: 1-696 <SCH>

A:Cross-references: EMBL:Z23261; NID:g113733; PIDN:CAA80786.1; PID:g113737

R:Dubois, E.; el Bakoury, M.; Glandsdorff, N.; Messenguy, F.; Pierard, A.; Schierens,

submitted to the Protein Sequence Database, August 1994

A:Reference number: S45782

A:Accession: S45796

A:Molecule type: DNA

A:Residues: 1-696 <DUB>

A:Cross-references: EMBL:Z35823; NID:g536095; PID:g536097; MIPS:YBL061C

R:Kawamoto, S.; Nomura, M.; Ohno, T.

J. Ferment. Bioeng. 74, 199-208, 1992

A:Title: Cloning and characterization of SKT5, a Saccharomyces cerevisiae gene that a

A:Accession: S45483

A:Reference number: S45483

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 15-532, 'RTGRRLISYAGTVD' <KAW>

C:Genetics:

A:Gene: SGD:SKT5

A:Cross-references: SGD:S0000157; MIPS:YBL061C

A:Map position: 2L

A:Keywords: calcium binding; EF hand; transmembrane protein

F:238-250/Domain: calcium binding #status predicted <CA1>

F:415-434/Domain: transmembrane #status predicted <TM1>

F:451-475/Domain: transmembrane #status predicted <TM2>

Query Match

Best Local Similarity

Matches 30; Conservative 28; Mismatches 51; Indels 13; Gaps 5;

QY 16 TIASCSTPESNPKNS---ANTLTSLIKHAVKOTCOTLTGHQYKIAMKLSSESKAK 72

Db 111 SLSSLGSTPTNSPPALROTNSSTLTKEQIKRRRSVDLSHW-----LTNGSSDQ 164

QY 73 ISFTACGCVADKPAEVS--LELTAAINPNAR--EVAOKIYRHSLSKPCMLETVNAFV 129

Db 165 LKATN-ESVADLSHQWISRYLGGKNTSLVPRKLTIEYRQNVKRSKDEPVLFOYQYWL 223

QY 130 PT 131

Db 224 QT 225

RESULT 32

S57180

probable membrane protein YJR151c - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein J2223; serine/chreonine-rich protein YJR151c

```

C:Species: Saccharomyces cerevisiae
C:Date: 23-Aug-1995 #sequence_revision 08-Sep-1995 #text_change 19-Apr-2002
C:Accession: S57180
R:Scarce, T.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S57169
A:Accession: S57180
A:Molecule type: DNA
A:Residues: 1-1161 <SCA>
A:Cross-references: EMBL:Z49651; NID:g1015902; PID:g1015903; GSPDB:GN00010; MIPS:YJR151c
C:Genetics:
A:Gene: SGD:DNA4; MIPS:YJR151c
A:Cross-references: SGD:S0003912
A:Map position: 10R
C:Keywords: transmembrane protein

Query Match 10.4%; Score 70.5; DB 2; Length 1161;
Best Local Similarity 18.8%; Pred. No. 1.1e+02;
Matches 24; Conservative 35; Mismatches 44; Indels 25; Gaps 5;

OY 4 ILVYATATMTAFTLASCASTPESNPKNSANLTTSLIKHAVKOTCOTOLTGHOYKRIAM 63
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 310 VISTATATSTPFASTLPATSTASTDHTTSSVST---NAFTSATTTTTSPTY----- 360

OY 64 KLSSSKAKISTACGCYADKAPKPEAVSLTELTTAANINAPKAEVAKIVRIKLAKCMLET 123
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 361 -ISSSPQSVSSA-----EPTTVSEV-TSSVEPTRSSQVTS-----SAEPTTVSE 404

OY 124 VNAPFVPT 131
      ::||
DB 405 FTSSVEPT 412

RESULT 33
T25697
hypothetical protein F16F9.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T25697
R:Fulton, B.
submitted to the EMBL Data Library, August 1996
A:Description: The sequence of C. elegans cosmid F16F9.
A:Reference number: Z20071
A:Accession: T25697
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1229 <FUT>
A:Cross-references: EMBL:U67956; PID:NAB07691.1; GSPDB:GN00028; CESP:F16F9.2
A:Experimental source: strain Bristol N2; clone F16F9
C:Genetics:
A:Gene: CESP:F16F9.2
A:Map position: X
A:introns: 35/1; 364/1; 482/2; 517/1; 971/1; 1021/2; 1179/3

Query Match 10.4%; Score 70.5; DB 2; Length 1229;
Best Local Similarity 23.4%; Pred. No. 1.2e+02;
Matches 32; Conservative 24; Mismatches 66; Indels 15; Gaps 4;

OY 8 TATMTAFTLASCASTPESNPKNSANLTTSLIKHAVKOTCOTOLTGHOYKRIAMKLS 67
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 800 TTTTTEKTSKTTTEKPTTSATTEETSTESTESTESTESTESTESTESTESTESTAE--TT 857

OY 68 ESKAKISET---ACGCYADKAPKPEAVSL-----TELTTAANINAPKAEVAKIVRIHSL 116
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 858 TTSATSETTTSESAAFTGESPEPTALOSSQKSEENSSAEKFGARDPVPKHKHTTV 917

OY 117 KPCMLETVNAPFVPT 133
      |||||
DB 918 KPA--ETTSVAASSTT 932

RESULT 34
A11280

```

[illegible]

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 16:01:02 ; Search time 13 Seconds

(without alignments)
427.525 Million cell updates/sec

Title: US-10-048-196-2

Perfect score: 676
Sequence: 1 MMKILYVATATMTAFTLASC.....SLKPCMLETVNAFIVPTTTR 134

Scoring table: BLISSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt.40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80.5	11.9	3083	1	POLG_ZYMVS
2	80	11.8	262	1	MI43_MYCIT
3	79.5	11.8	4377	1	ANK3_HUMAN
4	76.5	11.3	567	1	CH13_CANAL
5	76.5	11.3	605	1	YHC8_YEAST
6	76	11.2	995	1	Y109_YEAST
7	76	11.2	1041	1	EGT2_YEAST
8	74.5	11.0	495	1	YHJ2_YEAST
9	74	10.9	298	1	YHJ2_YEAST
10	73.5	10.9	1758	1	YHJ2_YEAST
11	73.5	10.9	1758	1	YHJ2_YEAST
12	72.5	10.7	928	1	RSC1_YEAST
13	72.5	10.7	1120	1	STFR_ECOLI
14	72.5	10.7	1140	1	YHJ2_YEAST
15	72.5	10.7	2715	1	G156_PARP
16	71	10.5	900	1	YHJ2_YEAST
17	70.5	10.4	676	1	YHJ2_YEAST
18	70.5	10.4	696	1	YHJ2_YEAST
19	70.5	10.4	1161	1	DAM4_YEAST
20	70	10.4	561	1	YHJ2_YEAST
21	70	10.4	562	1	YHJ2_YEAST
22	70	10.4	1260	1	ALSI_CANAL
23	69	10.2	910	1	DNM1_MYCPN
24	69	10.2	993	1	RPO1_ARATH
25	69	10.2	3083	1	POLG_ZYMVS
26	68.5	10.1	797	1	YHJ2_YEAST
27	68.5	10.1	1059	1	YHJ2_YEAST
28	68.5	10.1	1306	1	MSB2_YEAST
29	68	10.1	271	1	MURI_LACTA
30	67.5	10.0	184	1	PMFA_PROMI
31	67.5	10.0	646	1	IRG1_MOUSE
32	67.5	10.0	1169	1	YHJ2_YEAST
33	67.5	10.0	1165	1	MAPK_DROME

34	66.5	9.8	1092	1	NCA2_XENLA	P36335 xenopus lae
35	66	9.8	82	1	NXSB_NAJAT	P80958 naja atra (
36	66	9.8	191	1	APL2_PETMA	P07096 petromyzon
37	66	9.8	331	1	YB30_ARATH	O82314 arabidopsis
38	66	9.8	512	1	66D2_DROME	O94223 drosophila
39	66	9.8	526	1	CEAL_HUMAN	P13688 homo sapien
40	66	9.8	574	1	GAGJ_DROFU	P21331 drosophila
41	66	9.8	899	1	VI20_HSVSA	O01055 herpesvirus
42	66	9.8	1073	1	HR38_DROME	P49869 drosophila
43	66	9.8	1704	1	CB7D_CAEEL	P34358 caenorhabdi
44	65.5	9.7	1261	1	YDEI_SCHPO	O10451 schizosach
45	65.5	9.7	1356	1	HET1_PODAN	O00808 podospira a

ALIGNMENTS

RESULT 1
POLG_ZYMVS STANDARD: PRT: 3083 AA.

AC O36979:

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Genome polypeptide (Contains: N-terminal protein (P1); Helper

DE component proteinase (EC 3.4.22.45) (HC-Pro); Protein P3; 6 kDa

DE protein 1 (6K1); Cytoplasmic inclusion protein (CI); 6 kDa protein 2

DE (6K2); Genome-linked protein (VPG); Nuclear inclusion protein A (NI-A)

DE (NI-A) (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-Pro); Nuclear

DE inclusion protein B (NI-B) (NIB) (RNA-directed RNA polymerase)

DE (EC 2.7.7.48); Coat protein (CP)]

DE Zucchini yellow mosaic virus (strain Singapore) (ZYMV).

DE Viruses: ssRNA positive-strand viruses, no DNA stage; Polyviridae;

DE Polyvirus.

DE NCBI_TaxID=117130;

DE [1]

DE SEQUENCE FROM N.A.

DE Lee K.C., Wong S.M.;

DE Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DE TRANSMISSION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHD

DE FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT

DE MAY BE INVOLVED IN REPLICATION.

DE FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.

DE CATALYTIC ACTIVITY: Hydrolyzes glutamyl bonds, and activity is

DE further restricted by preferences for the amino acids in p6 'p1'

DE that vary with the species of polyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-

DE Glu+(Ser or Gly) for the enzyme from tobacco etch virus. The

DE natural substrate is the viral polyprotein, but other proteins and

DE oligopeptides containing the appropriate consensus sequence are

DE also cleaved.

DE CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +

DE (RNA)(N).

DE CATALYTIC ACTIVITY: Hydrolyzes a Gly-I-Gly bond at its own C-

DE terminus, commonly in the sequence -Tyr-Xaa-Val-Gly-I-Gly, in the

DE processing of the polyviral polyprotein.

DE PPM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.

DE PPM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE

DE POLYPEPTIDE WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC

DE PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT

DE INDIVIDUAL PROTEINS.

DE SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.

DE SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.

DE SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPEPTIDE FAMILY.

DE THIS SWISS-PROT entry is copyright. It is produced through a collaboration

DE between the Swiss Institute of Bioinformatics and the EMBL outstation -

DE the European Bioinformatics Institute. There are no restrictions on its

DE use by non-profit institutions as long as its content is in no way

DE modified and this statement is not removed. Usage by and for commercial

DE entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

DE or send an email to license@isb-sib.ch).


```

DR InterPro: IPR000488; Death.
DR InterPro: IPR000906; ZUS.
DR Pfam: PF00023; ank; 24.
DR Pfam: PF00531; death; 1.
DR Pfam: PF00791; ZUS; 1.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK; 21.
DR SMART: SM00005; DEATH; 1.
DR SMART: SM00218; ZUS; 1.
DR PROSITE: PS50088; ANK_REPEAT; 21.
DR PROSITE: PS50297; ANK_REPEAT; 1.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat.
FT REPEAT 73 102
FT REPEAT 106 135
FT REPEAT 139 168
FT REPEAT 172 201
FT REPEAT 203 230
FT REPEAT 234 263
FT REPEAT 267 296
FT REPEAT 300 329
FT REPEAT 333 362
FT REPEAT 366 395
FT REPEAT 399 428
FT REPEAT 432 461
FT REPEAT 465 494
FT REPEAT 498 527
FT REPEAT 531 560
FT REPEAT 564 593
FT REPEAT 597 626
FT REPEAT 630 659
FT REPEAT 663 692
FT REPEAT 696 725
FT REPEAT 729 758
FT REPEAT 762 791
FT REPEAT 795 825
FT DOMAIN 1519 1898
FT DOMAIN 4090 4174
SQ SEQUENCE 4377 AA; 480399 MM; FA2379E5576B684 CRC64;

Query Match 11.8%; Score 79.5; DB 1; Length 4377;
Best Local Similarity 26.7%; Pred. No. 28;
Matches 28; Conservative 18; Mismatches 52; Indels 7; Gaps 4;

QY 14 AFLTASCASTPESNP-KNSSANLTSLIKHAVKOTCOTLTGHQYKIAAMKLSSESKAK 72
DB 3900 ALFTSSCVDKSRIPVKNTPRD---NII--AVRKACATOKOG-QPEKGKAKQLPSKLPVK 3953

QY 73 ISETAGCGVADKAPKPEAVSLTELTTAIIINPARKTEVAKIVRHSIK 117
DB 3954 VRSTCVTTTTTATTTTTTTTTTCTVAKVRKSQLKEVCKHSIE 3998

RESULT 4
CHT3_CANAL
ID CHT3_CANAL STANDARD; PRT; 567 AA.
AC P40954;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Chitinase 3 precursor (EC 3.2.1.14).
GN CHT3.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10261;
RX MEDLINE=95223977; PubMed=7708682;
RA McCreath K.J., Specht C.A., Robbins P.W.;
RT "Molecular cloning and characterization of chitinase genes from
Candida albicans.";
```

```

RL Proc. Natl. Acad. Sci. U.S.A. 92:2544-2548(1995).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U15801; AAA68016.1; -.
DR HSPB: P23472; 2HVW.
DR InterPro: IPR001579; Chitinase_18/2.
DR InterPro: IPR001223; Glyco_hydro_18.
DR Pfam: PF00704; Glyco_hydro_18; 1.
DR PROSITE: PS01095; CHITINASE_18; 1.
KW Hydrolyase; Glycosidase; Chitin degradation; Signal; Glycoprotein;
KW Chitin-binding; Multigene family.
FT SIGNAL 1
FT CHAIN 17 567
FT ACT_SITE 157 157
FT CARBOHYD 159 159
FT DOMAIN 319 436
FT SEQUENCE 567 AA; 60060 MM; DD843126F65E22C2 CRC64;

Query Match 11.3%; Score 76.5; DB 1; Length 567;
Best Local Similarity 23.7%; Pred. No. 5.1;
Matches 31; Conservative 23; Mismatches 64; Indels 13; Gaps 4;

QY 8 TATLMTAFTLASCASTPESNPKNSSANLTSLIKHAVKOTCOTLTGHQYKIAAMKLS 67
DB 325 TTTTSTSTTSSSSSSSKTSSTSTSTSTSSSTSSSTSSSTSSSTSSSTSS 384

QY 68 ESKAKISETAGCGVADKAPKPEAVSLTELTTAIIINPARKTEVAKIVRHSIKCMLET--VN 125
DB 385 TTSSQISSTTS-----TAP--TSSTSLSSSTISSTISASISDPTTSVTSSETTPVPPSSLS 436

QY 126 AFIVP---TTT 133
DB 437 AITIPGDSSTT 447

RESULT 5
YHCB_YEAST
ID YHCB_YEAST STANDARD; PRT; 605 AA.
AC P38739;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Hypothetical 63.8 kDa protein in GUT1-RIM1 intergenic region
DE precursor.
GN YHLO28W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
RA Nhan M., Rifkin L., Riles L., St Peter H., Trevasakis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
```

```

RT      VIII."
RL      Science 265:2077-2082(1994).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; 011583; AAB65040.1; -
DR      SGI; S48940; S48940.
DR      SGD; S0001020; MSC4.
DR      InterPro; IPR002889; WSC.
DR      Pfam; PF01822; WSC; 1.
DR      SMART; SM00321; WSC; 1.
DR      Hypothetical protein; Transmembrane; Signal.
FT      CHAIN 1 26
FT      SIGNAL 1 26
FT      DOMAIN 116 317
FT      TRANSMEM 415 435
FT      SEQUENCE 605 AA; 63807 MW; C9DFGCBAA9553811 CRC64;
SQ
Query Match
Best Local Similarity 11.3%; Score 76.5; DB 1; Length 605;
Matches 30; Conservative 24; Mismatches 71; Indels 9; Gaps 3;
OY      7 VTATLMTAFTLASCASTPESN-----PKSSANLTSLSIKHAVKOTCOTQLTGHWKIA 61
DB      154 LSTSTPLTASTSTPSTPDTSTALPTTSTKLSIPTSTSTSTSTSS--TST 211
OY      62 AMKLSSEKAKISSTACGCVADKAPAVSLTELTT--AANPNAETVAOKIVRHSKLPK 119
DB      212 TVSVTSSTSTSTSTSTSTLSTSTSSSSSTPTTSSAPISSTSTSTSTSTPTSS 271
OY      120 MLETVNAFIVPTTT 133
DB      272 SAPTSSSTPTTST 285
DB
RESULT 6
YI09_YEAST
ID      YI09_YEAST STANDARD; PRT; 995 AA.
AC      P40442;
DT      01-FEB-1995 (Rel. 31, Created)
DT      01-FEB-1995 (Rel. 31, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Hypothetical 99.7 kDa protein in SDL1 5' region precursor.
GN      YIIL69C OR YI9402.07C.
OS      Saccharomyces cerevisiae (Baker's yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Saccharomycos.
OX      NCBI_TaxID=4932;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=S288c / AB972;
RA      Barrett B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA      Churche C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA      Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
RA      Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA      Rajadream K.A., Riles L., Rowley N., Skelton J., Smith V.,
RA      Walsh S.V., Whitehead S.;
RL      Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----

```

```

DR      EMBL; 246921; CAA87023.1; -
DR      SGD; S0001431; YIIL69C.
DR      InterPro; IPR004089; Chmtaxis_transd.
DR      InterPro; IPR000727; T-SNARE.
DR      PROSITE; PS50192; T-SNARE; UNKNOWN_1.
KW      Hypothetical protein; Signal.
FT      SIGNAL 1 23
FT      CHAIN 24 995
FT      DOMAIN 92 154
FT      CARBOHYD 28 28
FT      CARBOHYD 35 35
FT      CARBOHYD 468 468
FT      CARBOHYD 664 664
FT      SEQUENCE 995 AA; 99735 MW; F63E287A03F137EC CRC64;
SQ
Query Match
Best Local Similarity 11.2%; Score 76; DB 1; Length 995;
Matches 35; Conservative 24; Mismatches 59; Indels 52; Gaps 6;
OY      7 VTATLMTAFTLASCASTPESN-----NPKSSANLTSLSIKHAVKOTCOTQLT 53
DB      686 ITTVPCSSSTTFTTSCDETCHVSTGAVTVSVSSKSYTTATVTHCDNGCNTKVT 745
OY      54 GHQYKIAMKLSSEKAKISSTAC--GC-----VADKAPAVSLT----- 92
DB      746 SECSKETSATTPSPKSYTTVTVTHCDNGCNTKVTSEAPATTTTYSOSYTTATVTH 805
OY      93 -----ELTTAANPNAETVAOKIVRHSKLPKMLETVNA 126
DB      806 CDDNCKTKTIVTSEAPAVTTTVP--KTYTATVTVQCDNCGSKTKVTYS 853
DB
RESULT 7
EGT2_YEAST
ID      EGT2_YEAST STANDARD; PRT; 1041 AA.
AC      P42835;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      01-NOV-1997 (Rel. 35, Last annotation update)
DE      EGT2 protein precursor (Early G1 transcript 2).
GN      EGT2 OR YNL327W OR N0320.
OS      Saccharomyces cerevisiae (Baker's yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Saccharomycos.
OX      NCBI_TaxID=4932;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=S288c / FY1676;
RA      Maitahl M., Niclaud J.-M., Lavesque H., Gallardin C.;
RA      MEDLINE=95373280; PubMed=7645347;
RT      *Sequencing analysis of a 15.4 kb fragment of yeast chromosome XIV
RT      identifies the RPD3, PAS8 and KRE1 loci, five new open reading
RT      frames";
RL      yeast 11:567-572(1995).
RN      [2]
RP      SEQUENCE FROM N.A., AND FUNCTION.
RC      MEDLINE=96251274; PubMed=8668141;
RA      Kovacech B., Nasmyth K., Schuster T.;
RA      *EGT2 gene transcription is induced predominantly by swi5 in early
RA      G1";
RL      Mol. Cell. Biol. 16:3264-3274(1996).
CC      -1- FUNCTION: SEEMS TO BE INVOLVED IN THE CORRECT TIMING OF CELL
CC      SEPARATION AFTER CYTOKINESIS, AS SEPARATION OF MUTANT DAUGHTER
CC      CELLS IS DELAYED. COULD EITHER BE AN ENZYME NECESSARY FOR GLUCANS-
CC      DEGRADATION OF THE CELL WALL AT THE NECK REGION BETWEEN MOTHER AND
CC      DAUGHTER CELLS OR A REGULATORY PROTEIN CONTROLLING THIS METABOLIC
CC      STEP. EXCLUSIVELY EXPRESSED BETWEEN THE END OF MITOSIS AND EARLY
CC      G1; INACTIVATED BEFORE CELLS PASS START.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way

```

CC	modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC	-----
CC	EMBL; 246259; CAAB6371.1; -
DR	EMBL; 271603; CAAB6259.1; -
DR	SGD; 50005271; EGT2
KW	Glycoprotein; Repeat; Signal; Cell cycle; Cell division; Mitosis.
FT	SIGNAL 1 20
FT	CHAIN 21 1041
FT	DOMAIN 200 203
FT	DOMAIN 381 384
FT	DOMAIN 398 395
FT	DOMAIN 490 493
FT	DOMAIN 586 589
FT	REPEAT 457 492
FT	REPEAT 577 606
FT	REPEAT 613 647
FT	REPEAT 716 745
FT	REPEAT 773 802
FT	REPEAT 811 840
FT	REPEAT 849 886
FT	REPEAT 927 924
FT	REPEAT 925 962
FT	CARBOHD 65 65
FT	CARBOHD 103 103
FT	CARBOHD 161 161
FT	CARBOHD 175 175
FT	CARBOHD 249 249
FT	CARBOHD 332 332
FT	CARBOHD 401 401
FT	CARBOHD 435 435
FT	CARBOHD 465 465
FT	CARBOHD 485 485
FT	CARBOHD 506 506
FT	CARBOHD 526 526
FT	CARBOHD 544 544
FT	CARBOHD 556 556
FT	CARBOHD 635 635
FT	CARBOHD 636 636
FT	CARBOHD 657 657
FT	CARBOHD 709 709
FT	CARBOHD 756 756
SO	SEQUENCE 1041 AA; 108494 MW; 01FECCF8EAB744CD CRC64;
Query Match	11.2%; Score 76; DB 1; Length 1041;
Best Local Similarity	25.0%; Pred. No. 12;
Matches 30; Conservative 22; Mismatches 56; Indels 12; Gaps 3;	
OY	24 PESNKRSSANLTTSLLINHAVKOTCOTDLTGHOYK-----IAAMKLSSSEKAKISETA 77
DB	912 PVSITKSGKTSLSLST--BESVAVAGSTVYDASQVAEHNNIVNAIDPLKTSFQKATATEVC 968
OY	78 CGCVADKAPKAVSLTELTTLTAALNPAPKAEVAKIVRHS---LKPCMLETVNAFIYPTTTR 134
DB	969 VTCATLSSPHSATIDAGTTISLPTSSSTSLSTLIITWYSSSTIKPISISTVSGAAGQLTIR 1028
RESULT 8	
YHJ3_SALTY	
ID YHJ3_SALTY	STANDARD; PRT; 495 AA.
AC	P50335;
DT	01-OCT-1996 (Rel. 34, Created)
DT	15-JUN-2002 (Rel. 41, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Protein yHJ3 precursor.
GN	YHJ3 OR STM3613.
OS	Salmonella typhimurium.
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC	Salmonella.
OK	NCBI_TaxID=602;
RN	[1]
RP	SEQUENCE FROM N.A.

```

RC STRAIN=LT2 / SGC01412 / ATCC 700720;
RA MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., All J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.:
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RT Nature 413:852-856(2001).
RL [2]
RL SEQUENCE OF 1-287 FROM N.A.
RP
RC STRAIN=LT2;
RC MEDLINE=97113529; PubMed=8955389;
RA Baker K.E., Dittullio K.P., Neuhard J., Kella R.A.;
RA "Utilization of orotate as a pyrimidine source by Salmonella
RT typhimurium and Escherichia coli requires the dicarboxylate transport
RT protein encoded by dcfA.";
RT J. Bacteriol. 178:7099-7105(1996).
CC -1- SUBCELLULAR LOCATION: Periplasmic (Potential).
CC -1- MISCELLANEOUS: HAS LOST THE ACTIVE SITE RESIDUES.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: AE008867; AAL22473.1; -
DR EMBL: X91397; CA662743.1; ALT_INIT.
DR MEROPS: M16.UNA; -
DR Stygene: SG10743; VhJj.
DR InterPro: IPR001431; Peptidase_M16.
DR Pfam: PF00675; Peptidase_M16.
DR PROSITE: PS00143; INSULINASE; FALSE_NEG.
DR Hydrolase: Metalloprotease; Zinc; Periplasmic; Signal;
KW Complete proteome.
FT SIGNAL 1 24 BY SIMILARITY.
FT CHAIN 25 495 PROTEIN YHJj.
FT CONFLICT 244 244 G -> L (IN REF. 2).
FT CONFLICT 280 287 DTPQPIR -> IRGNRFA (IN REF. 2).
SQ SEQUENCE 495 AA; 55180 MM; 59D6A0198EB023B7 CRC64;

Query Match 11.0%; Score 74.5; DB 1; Length 495;
Best Local Similarity 21.9%; Pred. No. 6.9;
Matches 33; Conservative 21; Mismatches 58; Indels 39; Gaps 5;

OY 15 FTLASCASTPESNPKNSSANLTLSIRKIAVQOTCOTQLTGHWKIAMK----- 64
DB 331 FLRAQCAINIES--PNDRKLTNTLSLVANVELKAVRDKGSEEFNALVAKNLELOKLFAT 388
OY 65 -----LSSESKAKISEFYACGCAADKAPKAVS-----LTLETLTAIINPNAETVAOK 110
DB 389 YARTDTDLITGGRMSRLNQ---VVDIAPQOYOKLRONFLNSLTVDMLNOLRQOLSQE 444
OY 111 IYRHSIKP-----CMLETVNAFIYPTT 132
DB 445 MALILLOPQGEPEFNKALKATWDEIMVPTT 475

RESULT 9
VGLG_HRSVL
ID VGLG_HRSVL STANDARD; PRT; 298 AA.
AC P20895;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (subgroup A / strain Long).

```

OC Viruses: ssRNA negative-strand viruses; Mononegavirales;
 CC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OK NCBI_TaxID=11260;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA MEDLINE=87289657; PubMed=2441388;
 RX Johnson P.R., Springs M.K., Olmsted R.A., Collins P.L.;
 RT "The G glycoprotein of human respiratory syncytial viruses of
 RT antigenically related proteins."
 RL Proc. Natl. Acad. Sci. U.S.A. 84:5625-5629(1987).
 CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
 CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
 CC HEMAGGLUTININATING ACTIVITIES.
 CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
 CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
 CC -1- PPM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
 CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M1712; AAA47411.1; -
 DR PIR: A32703; MGZRL.
 DR InterPro: IPR000925; Glycoprotei_G.
 DR Pfam: PF00802; Glycoprotein_G; 1.
 KM Transmembrane; Glycoprotein.
 FT DOMAIN 1 37
 FT TRANSMEM 38 66
 FT DOMAIN 67 298
 FT CARBOHYD 103 135
 FT CARBOHYD 135 135
 FT CARBOHYD 237 237
 FT CARBOHYD 250 250
 FT CARBOHYD 251 251
 FT CARBOHYD 294 294
 SQ SEQUENCE 298 AA; 32781 MW; B79E9F4B4A73B0E CRC64;
 Query Match 10.9%; Score 74; DB 1; Length 298;
 Best Local Similarity 24.3%; Pred. No. 4.3;
 Matches 27; Conservative 20; Mismatches 50; Indels 14; Gaps 3;
 OY 1 MKKILVVTMTATFLASCASPESNPKNSSANLTSLIKHA---VKOTCOTLTGHQY 57
 DB 45 ILAMITSLIITAIITFIASA-----NHKVTLTATIIODATSIQIKNTTPYLTQDPQ 96
 OY 58 WKIAAMKLS---SESKAKISFTACGCVADKAPKAVSLTETLTAAINPNAKT 105
 DB 97 LGISFSNLSSEITSQTTTLASTTPGVKSNLOPTVYKTKKNTTTPOTOPSKPT 147
 RESULT 10
 YIR7_YEAST
 ID YIR7_YEAST STANDARD: PRT; 1758 AA.
 AC P40434;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 197.5 kDa protein in SDL1 5' region.
 GN Y1177C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes;
 OK NCBI_TaxID=4932;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Bartell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,

RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
 RA Gentles S., Hamlyn N., Hornsell T.S., Hunt S., Jagels K., Jones M.,
 RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
 RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
 RA Walsh S.V., Whitehead S.;
 RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. YEAST SUBTELOMERIC Y'
 CC REPEAT SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: Z46921; CA87015.1; -
 DR SGD: S0001439; Y1177C.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00271; helicase_C; 1.
 DR SMART: SM00490; HELIC_C; 1.
 KW Hypothetical protein; Hydrolyase; Helicase; ATP-binding; Repeat.
 FT NP_BIND 681 688
 FT SEQUENCE 1758 AA; 197511 MW; 9A191A3FEF7FBD7D CRC64;
 Query Match 10.9%; Score 73.5; DB 1; Length 1758;
 Best Local Similarity 26.0%; Pred. No. 37;
 Matches 33; Conservative 18; Mismatches 61; Indels 15; Gaps 5;
 OY 8 TATLTMTATFLASCASPESNPKNSSANLTSLIKHAVQCQQTQLTGHQYKIAAMKLS 67
 DB 1194 SATTASIVRTSATTESTEST--NSTNATTTTESTNSTNATTTTESTIN-----SMTSA 1243
 OY 68 ESKAKISFTACGCVADKAPKAVSLTETLTAAINPNAKT--VAKIVHSILKPCMLTETVNA 126
 DB 1244 TTASIVRTSATTESTESTNSTAT--TTAST--NVRTSATTTTSINSNSTNATTTTESTINS 1299
 OY 127 FIVPTTT 133
 DB 1300 NTNATTT 1306
 RESULT 11
 YJW5_YEAST
 ID YJW5_YEAST STANDARD: PRT; 1758 AA.
 AC P40889; Q05371; Q9UQW2;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 197.6 kDa protein in FSP2 5' region.
 GN YJ1225C OR J0202 OR HR393/HRD732 OR HR393/HRD1054.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes;
 OK NCBI_TaxID=4932;
 RN (1)
 RP PRELIMINARY SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA MEDLINE=95242842; PubMed=7725802;
 RA Vandenbol M., Durand P., Bolle P.-A., Dion C., Portetelle D.,
 RA Hilger F.;
 RT "Sequence analysis of a 40.2 kb DNA fragment located near the left
 RT telomere of yeast chromosome X.";
 RL Yeast 10:1657-1662(1994).
 RN (2)
 RP REVISIONS.
 RA Vandenbol M., Durand P., Portetelle D., Hilger F.;
 RL Submitted (SEP-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. YEAST SUBTELOMERIC Y'
 CC REPEAT SUBFAMILY.

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

DR EMBL; 249498; CAA89520.1; -
DR EMBL; 249500; CAA89522.1; -
DR EMBL; 248148; CAA88141.1; ALT_SEQ.
DR EMBL; 248148; CAA88142.1; ALT_SEQ.
DR EMBL; 234098; CAA83986.1; ALT_SEQ.
DR EMBL; 234098; CAA83985.1; ALT_SEQ.
DR PIR; S45148; S45148.
DR PIR; S45147; S45147.
DR SGD; S0003760; YJ1225C.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00490; HelicC; 1.
KM Hypothetical protein; Hydrolyase; Helicase; ATP-binding; Repeat.
FT NP_BIND 681 688 ATP (POTENTIAL).
SQ SEQUENCE 1758 AA; 197564 MW; 7E1D052AF97F8CA0 CRC64;

Query Match 10.9%; Score 73.5; DB 1; Length 1758;
Best Local Similarity 26.0%; Pred. No. 37;
Matches 33; Conservative 18; Mismatches 61; Indels 15; Gaps 5;

QY 8 TATLMAFTLASCAPSPESNPKSSANLTTSLIKHAVKQTCOTQLNGHWKIAMKLSS 67
DB 1194 SATTASINVRTSATTTSTSTSTAT--NSSTNATTTSTSTNATTTSTN-----SMTSA 1243
QY 68 ESRAKSEFACGVADKAPAVSITELTTAINDNATE--VAQKIYRHSLKPCMLETVA 126
DB 1244 TTTASINVRTSATTTSTSTSTAT--TTASI--NVRTSATTTKTSINSSTNATTTSTNS 1299
QY 127 FIVPTT 133
DB 1300 NTNATTT 1306

RESULT 12
RSC1_YEAST
ID RSC1_YEAST STANDARD; PRT; 928 AA.
AC P53236;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chromatin structure remodeling complex protein RSC1.
GN RSC1 OR YGR056W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OX Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Eutlan K.D., Rose M., Koetter P., Roehmer A., Schram I.,
RA Hempel S.;
RL Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: COMPONENT OF THE RSC CHROMATIN REMODELING COMPLEX. RSC
CC IS RESPONSIBLE FOR THE TRANSFER OF A HISTONE OCTAMER FROM A
CC NUCLEOSOME CORE PARTICLE TO NAKED DNA.
CC -1- SUBUNIT: RSC IS COMPOSED OF 15 SUBUNITS; AMONG WHICH ARP7, ARP9,
CC RSC1, RSC2, RSC4, RSC6, RSC8, SFH1 AND STH1.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: CONTAINS 1 BAH DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 BROMODOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

DR EMBL; 272841; CAA97057.1; -
DR HSSP; Q92831; IB91.
DR SGD; S0003288; RSC1.
DR InterPro; IPR001025; BAH.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR Pfam; PF01426; BAH; 1.
DR SMART; SM00439; BAH; 1.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 1.
DR PROSITE; PS00633; BROMODOMAIN_2; 2.
DR PROSITE; PS00633; BROMODOMAIN_2; 2.
KM Chromatin regulator; Nuclear protein; Bromodomain; Repeat.
FT DOMAIN 27 95 BROMODOMAIN 1.
FT DOMAIN 255 325 BROMODOMAIN 2.
SQ SEQUENCE 928 AA; 106669 MW; EFB0922FC08EC27 CRC64;

Query Match 10.7%; Score 72.5; DB 1; Length 928;
Best Local Similarity 24.2%; Pred. No. 22;
Matches 39; Conservative 22; Mismatches 59; Indels 41; Gaps 7;

QY 3 KILVTVATIMTAF-----TLASCAPSPESNPKSS-----ANLT 36
DB 584 KIDETGTITITDITLTSSMPRVNSSSTRIPPTLKQTSIPSSNRSSNPPLQHONQOT 643
QY 37 TSLIKHAVKQTCOTQLNGH-----QYWKIAMKLSSSEK-----AKISFACGVADKAPAV 89
DB 644 SNFLKLENNMNSHNLSHSVKPFQSPSLLEGSRKRYSAK KQQLSTARKKASK 702
QY 90 SLEELTTAINDNATEVAQKIYRHSLKPCMLETVA 130
DB 703 SFT-ISSMINTLTAHTS-----KYNFNHIVTEAPGAFVVP 736

RESULT 13
STR_ECOLI
ID STR_ECOLI STANDARD; PRT; 1120 AA.
AC P76072; P77560;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Side tail fiber protein homolog from lambdaoid prophage Rac.
GN STRF OR B1372.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; Pubmed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT *The complete genome sequence of Escherichia coli K-12.*;
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; Pubmed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa H.,
RA Makino K., Miki T., Mizoduchi K., Mori H., Mori T., Motomura K.,
RA Nakae S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA Sampaio G., Seki Y., Sivasubaram S., Tagami H., Takeda J.,
RA Takeuchi K., Takeuchi Y., Wada C., Yamamoto Y., Horikuchi T.,
RT *A 570-kb DNA sequence of the Escherichia coli K-12 genome

```
RF      corresponding to the 28-0-40.1 min region on the linkage map.";
RL DNA Res. 3;383-377(1996).
CC -I SIMILARITY: BELONGS TO THE TAIL FIBER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation .
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as their consent is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licensese@isb-sib.ch)).
CC -----
DR EMBL; AE000234; AAC7A454.1; ALT_INIT.
DR DDB; D90774; BAI1966.1; -.
DR EMBL; D90775; BAI1975.1; -.
DR Ecogene; EGJ3370; sStr.
DR InterPro: IPR004089; ChmCaxis-transd.
DR InterPro: IPR005003; Phage_fiber_
DR InterPro: IPR005068; Phage_fiber_2.
DR Pfam; PF03335; Phage_fiber_6.
DR Pfam; PF03406; Phage_fiber_2; 1.
KW Hypothetical protein; Fiber protein; Repeat; Complete proteome.
SQ SEQUENCE 1120 AA; 113779 MW; 542E5D71EEF95B4 CRC64;

Query Match          10 %; Score 72.5; DB 1; Length 1120;
Best Local Similarity 25.4%; Pred. No. 28;
Matches   35; Conservative    16; Mismatches    60; Indels     27; Gaps       4;

QY      8 TATLTMTAFTLASCASAPESN-----PKNSSANLTSLIKHAYKQTOTOLT 53
Db      212 SASLSQAATASTATTAKSEALNTSAADAASAASKASTTNASSASSAASTA---A 267
        |::|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      54 GHGYMKIAAMKLSSSEKAISTPAG----CVADKAEPAVSLTELTTTAIPNPARTFA 108
Db      268 GNS----AAAARKSETNNARSSETAAGQSASAAGA SKTAAASASASAVTSACGASASATA 323
        ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      109 QKIVRHSLLKPCLMEFYNA 126
Db      324 GKSAESAASASTATTTKA 341
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 14
YM96_YEAST STANDARD: PRT; 1140 AA.
ID YM96_YEAST O04893:
AC O04893:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypotheical 113.1 kDa protein in PRE5-Fer1 Intergenic region.
GN YMR317W OR YM992A.09.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycas.
OX NCBI_TaxID=4932; [1]
RN NCBIMSeq FROM N.A.
RP STRAIN=S288c / AB972;
RA Churcher C.M., Louis E.J., Barrell B.G., Rajandream M.A., Walsh S.V.;
RC Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
CU -I- DOMAIN: CONTAINS MANY SER/THR-RICH DOMAIN AND REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboratio
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (See http://www.isb-sib.ch/announc
CC or send an email to licensese@sib-slb.ch)).
CC -----
EMBL: Z54141; CA90835.1; -.
SCD; S0004936; YMR317W.
KW Hypothetical protein; Repeat.
SQ SEQUENCE 1140 AA; 113070 MW; 0153BCA24FE5427 CRC64;
```

```

Query Match      10.7%: Score 72.5: DB 1: Length 1140:
Best LocalSimilarity 23.1%: Pred. No. 28:
Matches 34: Conservative 28: Mismatches 54: Indels 31: Gaps 5:

OY 9 ATLMTAFTLASCASCPESNPKNSSANLTTSLIKHAVQTCOTOLGHQVYKIAMLISE 68
Db 372 IPLATSSVSEADAPSTSSSVSSAPSTSS---SSVSEAPSTSSSVSEISSTL 427
OY 69 SKATSEFACCCVADKAPAVS-----LTELTL--TAALNPARE 106
Db 428 MSSEVSSATSSVSEAPSAISSLASSRLFPSSKNTVSTSLVATEASSVTSLSRPSETL 487
OY 107 VAOKIVRSLKPCMLLEVNAFEVPTTT 133
Db 488 ASNSTIESTSTGYNSTVS-----TTTT 509

RESULT 15
G156_PARP
ID G156_PARP STANDARD; PRT; 2715 AA.
AC P13837;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-AUG-1990 (Rel. 15, Last annotation update)
DE 156G surface protein precursor.
GN 156G
OS Paramecium primaurelia.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Periculida;
OC Paramecium.
OX NCBI_TaxID=5886;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=156;
RX MEDLINE=87060934; PubMed=3783679;
RA Prat A., Kalinka M., Caron F., Meyer E.;
RT "Nucleotide sequence of the Paramecium primaurelia G surface protein.
RL J. Mol. Biol. 189:47-60(1986).
CC -!- FUNCTION: THIS PROTEIN IS THE SURFACE ANTIGEN OR IMMOBILIZATION
CC ANTIGEN OF PARAMECIUM PRIMAURELIA.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC MISCELLANEOUS: IT HAS INTERNAL HOMOLOGIES AND A HIGHLY PERIODIC
CC STRUCTURE WITH 37 PERIODS OF ABOUT 75 RESIDUES, EACH PERIOD
CC CONTAINING 8 CYSTEINES, EXCEPT FOR FOUR HALF PERIODS. A VARIABLE
CC PART OF 4/5 RESIDUES COMPRISSES 4 ALMOST IDENTICAL PERIODS IN THE
CC MIDDLE OF THE PROTEIN.
CC -!- MISCELLANEOUS: EXPRESSION OF G PROTEIN OCCURS AT LOW TEMPERATURES
CC (14-32 DEGREES CELSIUS).
CC -!- SIMILARITY: 98% TO THE ALLELIC FORM 168G PROTEIN (AC P17053) IN
CC PERIODIC STRUCTURE AND 80% IN VARIABLE DOMAIN IN THE MIDDLE OF
CC THE PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X03882; CAA27514.1; -.
DR DR
DR InterPro: IPR002895; Paramecium_SA.
DR Pfam: PF01508; Paramecium_SA; 34.
DR KMW Signal; Repeat; Antigen; Membrane; GPI-anchor.
DR FT SIGNAL 1 20 POTENTIAL.
DR FT CHAIN 21 2715 156G SURFACE PROTEIN.
DR FT DOMAIN 106 2560 37 x 75 AA APPROXIMATE REPEATS.
DR FT SIMILAR 1 222 88% TO PARAMECIUM TERRAURELIA A
DR FT PROTEIN.
DR SEQUENCE 2715 AA: 279551 MM: 97BE59AB9C7C298 CRC64:

```

[illegible]

OX NCBI_TaxID=4932;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C;
 RX MEDLINE=94205266; PubMed=8154187;
 RA Scherens B., el Bakkoury M., Vierendeels F., Dubois E., Messenguy F.;
 RT "Sequencing and functional analysis of a 32,560 bp segment on the
 RL left arm of yeast chromosome II. Identification of 26 open reading
 frames, including the KIP1 and SEC17 genes.";
 RN Yeast 9:1355-1371(1993).
 [2]
 RP SEQUENCE OF 1-678 FROM N.A.
 RC STRAIN-KR13;
 RX MEDLINE=93379362; PubMed=7764021;
 RA Kawamoko S., Sasaki T., Itahashi S., Hatsuyama Y., Ohno T.;
 RT "A mutant allele skt5 affecting protoplast regeneration and killer
 RL toxin resistance has double mutations in its wild-type structural
 gene in *Saccharomyces cerevisiae*.";
 RN Biosci. Biotechnol. Biochem. 57:1391-1393(1993).
 [3]
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
 RA Kawamoko S., Nomura M., Ohno T.;
 RT "Cloning and characterization of SKT5, a *Saccharomyces cerevisiae*
 RL toxin that affects protoplast regeneration and resistance to killer
 toxin of *Kluyveromyces fragilis*.";
 RN J. Ferment. Bioeng. 74:199-208(1992).
 [4]
 RP SUBUNIT, AND SUBCELLULAR LOCATION.
 RX MEDLINE=97461567; PubMed=9314530;
 RA Demarini D.J., Adams A.E., Fares H., De Virgilio C., Valle G.,
 RT Chuang J.S., Pringle J.R.;
 RT "A septin-based hierarchy of proteins required for localized
 RL deposition of chitin in the *Saccharomyces cerevisiae* cell wall.";
 CC J. Cell Biol. 139:75-93(1997).
 CC -1- FUNCTION: POSSIBLE ROLE IN PROTOPLAST REGENERATION AND KILLER
 CC TOXIN OF *K. LACTIS* (PGKL) RESISTANCE.
 CC -1- SUBUNIT: MAY INTERACT WITH CHS3 AND SEEMS TO BE AN ADAPTOR (ALONG
 CC WITH BNI4) TO LINK CHS3 TO SEPTINS.
 CC -1- SIMILARITY: SOME, TO S.POMBE SPAC24B11.10C AND TO YEAST YER096W.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-15 IS THE INITIATOR.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: Z23261; CA80786.1; -
 CC EMBL: Z35823; CA84882.1; -
 CC EMBL: S65415; AAC60564.1; ALT_INIT.
 CC PIR: S39827; S39827.
 CC PIR: S37328; S37328.
 CC SGD: S0000157; SKT5.
 CC InterPro: IPR001230; Prenyl_site.
 CC Prenylation; Lipoprotein.
 CC KW Lipid
 CC FT LIPID 693 FARNESYL (BY SIMILARITY).
 CC FT CONFLICT 350 Q -> E (IN REF. 2).
 CC FT CONFLICT 643 G -> T (IN REF. 2).
 CC SQ SEQUENCE 696 AA; 77066 MW; D50DE825E175D165 CRC64;
 Query Match 10.4%; Score 70.5; DB 1; Length 696;
 Best Local Similarity 24.6%; Pred. No. 25;
 Matches 30; Conservative 28; Mismatches 51; Indels 13; Gaps 5;
 OY 16 TLASGASPEENPKNS---ANLTSLIKHAVKOTQOTLTGHQYWKIAMKLSSESXAK 72
 DB 111 SLSSLGSTPTNPSFGALRQTNSSITLKEQIKRTRSVDSHMK-----LINSSTQ 164
 OY 73 ISETACGCVADKAPKAEVAVS-LTELTTAIPNART-EVAQKIVRHSKPCMLETVNATIV 129
 DB 165 LTAIIN-ESVADLSHOMISRIYLGKNNNTSLVPRKLTIEYRONVKKSKDPYELVFOYQVNL 223

OY 130 PT 131
 DB 224 QT 225
 RESULT 19
 DAN4_YEAST
 ID DAN4_YEAST STANDARD; PRT: 1161 AA.
 AC P47179;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cell wall protein DAN4 precursor.
 GN DAN4 OR YR151C OR YJ223.
 OS *Saccharomyces cerevisiae* (baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Scarcez T.;
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP REGULATION.
 RX MEDLINE=21113168; PubMed=11160904;
 RA Cohen B.D., Serfil O., Abramova N.E., Davies K.J., Lowry C.V.;
 RT "Induction and repression of DAN1 and the family of anaerobic
 RL mannoprotein genes in *Saccharomyces cerevisiae* occurs through a
 RT complex array of regulatory sites.";
 CC Nucleic Acids Res. 29:799-808(2001).
 CC -1- FUNCTION: COMPONENT OF THE CELL WALL (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC (potential).
 CC -1- PTM: EXTENSIVELY O-GLYCOSYLATED (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE SRP1 / TIPI FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: Z49651; CA89684.1; -
 CC SGD: S0003912; DAN4.
 CC InterPro: IPR000992; SRP1_TIP1.
 CC DR Pfam: PF00660; SRP1_TIP1.1.
 CC DR PROSITE: PS00724; SRP1_TIP1.1.
 CC KW Cell wall; Glycoprotein; Membrane; GPI-anchor; Signal.
 CC FT SIGNAL 1 24
 CC FT CHAIN 25 1146 CELL WALL PROTEIN DAN4.
 CC FT PROPEP 1147 1161 REMOVED IN MATURE FORM (POTENTIAL).
 CC FT LIPID 1146 1146 GPI-ANCHOR (POTENTIAL).
 CC SQ SEQUENCE 1161 AA; 118358 MW; 7954C1506970CA58 CRC64;
 Query Match 10.4%; Score 70.5; DB 1; Length 1161;
 Best Local Similarity 18.8%; Pred. No. 45;
 Matches 24; Conservative 35; Mismatches 44; Indels 25; Gaps 5;
 OY 4 ILVYATLTMTATTLASGASPEENPKNSANLTSLIKHAVKOTQOTLTGHQYWKIAM 63
 DB 310 VISTATITSTTASLTTPATSTASTDHTTSVSTY---NAFTSATTTTSDTY----- 360
 OY 64 KLSSEKAKISSETACGCVADKAPKAEVAVS-LTELTTAIPNART-EVAQKIVRHSKPCMLET 123
 DB 361 -ISSSSPOQVYSSA-----EPTIVSEV-TSSVEPTRSSQVYS-----SAEPTVSE 404
 OY 124 VNAFTVPT 131
 DB 405 FTSSVEPT 412

[illegible]

FT	TRANSMEM	375	399	7 (POTENTIAL)
FT	DOMAIN	400	561	CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	60	60	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	76	76	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	LIPID	413	413	PALMITATE (POTENTIAL).
FT	DOMAIN	21	57	GLY-RICH.
FT	DOMAIN	416	421	POLY-ARG.
FT	CONFLICT	424	483	AYGHHHNRASGTGARSDCAPSPRIAPPCAPLIPAHGAGSS
FT				APPPEODSSSSSRKPPASA -> SLRPLASLDRRAFRLR
FT				PPSHKSPRGSSPHCTPGCGLGRHAGDAGFGLDOOSKASLRL
FT				(IN REF. 1).
SO	SEQUENCE	561 AA;	59354 MW;	B6537DCAD4F7BE27 CRC64;
QY	Query Match	10.4%;	Score 70;	DB 1; Length 561;
Db	Best Local Similarity	26.1%;	Pred. No. 22;	
	Matches	29; Conservative	19; Mismatches	43; Indels 20; Gaps
QY	12	MTATFLASCATSEPNKNSANMLTSLIKH-----AVKQTCOTLTGHQYWKIAAKLSS	67	
Db	456	LTAHPGAGSADPTPTQDSVSSSKRPASALREMLGLPRLPTQLRA-----KVSSLISKI	511	
QY	68	ESKAKISETAGCGVADAPKAPASVLTETLTATAINPNARTEVAOKIVRHSIKP	118	
Db	512	RSGARRAETACALRSE--VEAVSL-----NVPOGAEAVLCQAYER	550	
RESULT 21				
CHIT_YEAST	STANDARD;			PRT; 562 AA.
ID	CHIT_YEAST	STANDARD;		PRT; 562 AA.
AC	P29029; P29028;			
DT	01-DEC-1992 (Rel. 24, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Endochitinase precursor (EC 3.2.1.14) (Soluble cell wall protein 2).			
GN	CTSI OR SCMW OR YLR286C OR L8003.13.			
OS	Saccharomyces cerevisiae (Baker's Yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.			
OX	NCBI_Taxid:4932;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 21-38.			
RC	STRAIN-DBY918, AND DBY939;			
RC	STRAIN-DBY918, AND DBY939;			
RX	MEDLINE-92011782; PubMed-1918080;			
RA	Kuranda M.J., Robbins P.W.:			
RA	"Chitinase is required for cell separation during growth of			
RA	Saccharomyces cerevisiae."			
RT	J. Biol. Chem. 266:19758-19767(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-S288C / AB972;			
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,			
RA	Johnston A., Fulton L., Galling S., Greco T., Kirsten J.,			
RA	Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,			
RA	Johnson D., Johnston L., Langston Y., Latreille P., Le T.,			
RA	Mardis E., Meneses S., Miller N., Nhan M., Pauley A., Peluso D.,			
RA	Rikken L., Riles L., Tatch A., Trevasis E., Vignati D.,			
RA	Wilcox L., Wohldman P., Vaudin M., Wilson R., Waterston R.;			
RA	Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.			
RN	[3]			
RP	SEQUENCE OF 21-32, AND SUBCELLULAR LOCATION.			
RC	STRAIN-SE6210;			
RC	MEDLINE-98422453; PubMed-9748433;			
RA	Cappellaro C., Mrsa V., Tanner W.:			
RA	"New potential cell wall glucanases of Saccharomyces cerevisiae and			
RT	their involvement in mating."			
RL	J. Bacteriol. 180:5030-5037(1998).			
CC	-1- FUNCTION: CHITINASE IS REQUIRED FOR CELL SEPARATION DURING GROWTH			
CC	OF SACCCHAROMYCES CEREVISIAE.			
CC	-1- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-			
CC	acetyl-D-glucosamine polymers of chitin.			
CC	-1- SUBCELLULAR LOCATION: MOST OF THE ENZYME IS SECRETED, BUT A			
CC	SIGNIFICANT AMOUNT OF CHITINASE IS ALSO FOUND ASSOCIATED WITH THE			
CC	CELL WALL THROUGH BINDING OF C-TERMINAL DOMAIN TO CHITIN.			

CC -1- PTM: EXTENSIVELY GLYCOSYLATED WITH A SERIES OF SHORT O-LINKED
CC MANNOSE OLIGOSACCHARIDES RANGING IN SIZE FROM MAN(2) TO MAN(5).
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M74070; AAA34539.1; -
DR EMBL: M74069; AAA34538.1; -
DR EMBL: U17243; AAB67331.1; -
DR PIR: B41035; B41035.
DR PIR: A41035; A41035.
DR HSSP: P23472; 2HYM.
DR SGD: S0004276; CTS1.
DR InterPro: IPR005089; CBM_19.
DR InterPro: IPR001579; Chitinase_18/2.
DR InterPro: IPR001223; Glyco_hydro_18.
DR Pfam: PF00704; Glyco_hydro_18; 1.
DR Pfam: PF03427; CBM_19; 1.
DR PROSITE: PS01095; CHITINASE_18; 1.
DR HydroLase: Glycosidase; Chitin degradation; Chitin-binding;
KW Cell wall; Signal; Glycoprotein.
FT SIGNAL 1 20
FT CHAIN 21 562
FT DOMAIN 21 327
FT DOMAIN 328 480
FT DOMAIN 481 562
FT ACT_SITE 157 157
FT VARIANT 16 16
FT VARIANT 23 23
FT VARIANT 321 321
FT VARIANT 336 340
FT VARIANT 399 399
FT VARIANT 433 434
FT VARIANT 461 461
FT VARIANT 477 481
FT CONFLICT 168 168
FT SEQUENCE 562 AA; 59014 MW; 0ABCEFB44BE1E19 CRC64;
Query Match 10.4%; Score 70; DB 1; Length 562;
Best Local Similarity 22.3%; Pred. No. 22;
Matches 37; Conservative 26; Mismatches 49; Indels 54; Gaps 7;
OY 8 TATLTMTAFTLASCASCTPESNPKNSANLTT-----SLIKHAWKQC 48
DB 316 TTTVAITSKTSASTSSASTSASTSOKTKTQSTTSQSKSVTLSPPTASAITSTIOTT 375
OY 49 QTQLTGHQYKIAAMKLSSESKAKIS-----ETACGCYADAKPEAVSLTELTTAA-INP 101
DB 376 KT-----LISSTKTSLSLGTSTTESLNSVAITSMKTTLSQSALTVP 421
OY 102 NART-----EVAQKIVRHSKPC-----MLET-VNAFIVPTTT 133
DB 422 QTTSIVSSAPIQTAITSTLSPTARKSSSVSLQTATTSTLSPPTT 467
RESULT 22
ALSL_CANAL STANDARD: PRT; 1260 AA.
AC P46590;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Agglutinin-like protein 1 precursor.
GN ALS1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC -----
CC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-ATCC 11651 / B792;
RX MEDLINE=95272392; PubMed=7752895;
RA Hoyer L.L., Scherer S., Shatzman A.R., Livi G.P.;
RT "Candida albicans ALS1: domains related to a Saccharomyces cerevisiae
RL sexual agglutinin separated by a repeating motif."
RL Mol. Microbiol. 15:39-54(1995).
CC -1- FUNCTION: MAY PLAY A ROLE IN ADHESION AND PATHOGENESIS.
CC -1- PTM: N-GLYCOSYLATED AND O-GLYCOSYLATED (POTENTIAL).
CC -1- SIMILARITY: TO YEAST SAG1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L25902; AAC41649.2; -
DR Cell adhesion; Glycoprotein;
KW Cell adhesion; Glycoprotein;
FT SIGNAL 1 17
FT CHAIN 18 1260
FT DOMAIN 1 1260
FT REPEAT 433 792
FT REPEAT 433 468
FT REPEAT 469 504
FT REPEAT 505 540
FT REPEAT 541 576
FT REPEAT 577 612
FT REPEAT 613 648
FT REPEAT 649 684
FT REPEAT 685 720
FT REPEAT 721 736
FT REPEAT 757 792
FT DOMAIN 983 1152
FT REPEAT 983 1043
FT REPEAT 1092 1152
FT DOMAIN 399 404
FT DOMAIN 408 418
FT DOMAIN 450 455
FT DOMAIN 486 491
FT DOMAIN 522 527
FT DOMAIN 558 563
FT DOMAIN 594 599
FT DOMAIN 630 635
FT DOMAIN 666 671
FT DOMAIN 702 707
FT DOMAIN 738 743
FT DOMAIN 774 779
FT DOMAIN 874 877
FT CARBOHYD 471 471
FT CARBOHYD 471 471
FT CARBOHYD 579 579
FT CARBOHYD 615 615
FT CARBOHYD 687 687
FT CARBOHYD 723 723
FT CARBOHYD 723 723
FT CARBOHYD 820 820
FT CARBOHYD 886 886
FT CARBOHYD 918 918
FT CARBOHYD 973 973
FT CARBOHYD 1045 1045
FT CARBOHYD 1068 1068
FT SEQUENCE 1260 AA; 132641 MW; 763D1063A2354C24 CRC64;
Query Match 10.4%; Score 70; DB 1; Length 1260;
Best Local Similarity 23.6%; Pred. No. 56;
Matches 33; Conservative 27; Mismatches 54; Indels 26; Gaps 4;
OY 7 VTATLTMTAFTLASCASCTPESNPKNSANLTTSLIKHAWKQCQTQLTGHQ----- 56
DB 891 VTSTVPTASTWSDSLSSYDGISATSSDNVS-----KSGSVSTTETTSVTITQTTPNPLSSV 946

[illegible]

```

1  15-DEC-1998 (Rel. 37, Created)
2  15-DEC-1998 (Rel. 37, Last sequence update)
3  16-OCT-2001 (Rel. 40, Last annotation update)
4  DNA-directed RNA polymerase, chloroplast precursor (EC 2.7.7.6).
5  RPOPT OR A2G24120 OR F27D4.3.
6  Arabidopsis thaliana (mouse-ear cress).
7  Eukaryota; Viridiplantae; Streptophyta;
8  Charophyta; Magnoliophyta; eudicotyledons: core eudicots; Rosidae;
9  eucosids II; Brassicales; Brassicaceae; Arabidopsis.
10 NCBI_TaxId=3702;
11 [1]
12 SEQUENCE FROM N.A.
13 STRAIN=cv. Columbia;
14 MEDLINE=97390506; PubMed=9242608;
15 Hedtke B., Boerner T., Weihe A.;
16 "Mitochondrial and chloroplast phage-type RNA polymerases in
17 Arabidopsis."
18 Science 277:809-811(1997).
19 [2]
20 SEQUENCE FROM N.A.
21 STRAIN=cv. Columbia;
22 MEDLINE=20083487; PubMed=10617197;
23 Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
24 Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldbaum T.V.,
25 Buell C.S., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
26 Moffat K.S., Cronin L.A., Shen M., Pei G., Van Aken S., Umayam L.,
27 Tallon L.J., Gill J.E., Adams M.D., Cartera A.J., Creasy T.H.,
28 Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
29 Nieman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
30 Venter J.C.;
31 "Sequence and analysis of chromosome 2 of the plant Arabidopsis
32 thaliana."
33 Nature 402:761-768(1999).
34 RL
35 -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
36 OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
37 SUBSTRATES.
38 CC
39 -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
40 [RNA](N).
41 CC
42 -1- SUBCELLULAR LOCATION: Chloroplast.
43 CC
44 -1- SIMILARITY: BELONGS TO THE PHAGE AND MITOCHONDRIAL RNA POLYMERASES
45 FAMILY.
46 -----
47 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
48 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
49 CC the European Bioinformatics Institute. There are no restrictions on its
50 CC use by non-profit institutions as long as its content is in no way
51 CC modified and this statement is not removed. Usage by and for commercial
52 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
53 CC or send an email to license@sib-sib.ch).
54 CC -----
55 DR EMBL; Y08722; CAA69972.1; -
56 DR EMBL; Y08463; CAA69717.1; -
57 DR EMBL; AC005967; AAD03373.1; -
58 HSP, P00573; IARO.
59 DR InterPro; IPR002092; RNA_pol_phage.
60 DR Pfam; PF00940; RNA_pol_1.
61 DR PROSITE; PS00900; RNA_POL_PHAGE.1; 1.
62 DR PROSITE; PS00489; RNA_POL_PHAGE.2; 1.
63 KW Transferase; DNA-directed RNA polymerase; Transcription;
64 KM Chloroplast; Transl peptide.
65 FT TRANSIT 1 ?
66 FT CHAIN ? 993
67 FT ACT_SITE 694 694 BY SIMILARITY.
68 FT ACT_SITE 769 769 BY SIMILARITY.
69 FT ACT_SITE 926 926 BY SIMILARITY.
70 SO SEQUENCE 993 AA; 112623 MW; BFBBC748DEB1445D CRC64;
71
72 Query Match 10.2%; Score 69; DB 1; Length 993;
73 Best Local Similarity 24.8%; Pred. No. 53;
74 Matches 34; Conservative 19; Mismatches 42; Indels 42; Gaps 6;
75
76 24 PESNPKSSAN-LFTSLIKHAVKOTCOTOLTGHOYWKIA--KLSLSERK----- 70
77 ||||| : : : ||||| : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db 747 PESNPALAKILITOVDRUKOTVWTSYGVY--VGAREQIKRRLKEKGVITDERML 804
 QY 71 -----AKISSTAGCGVADKAPKAPV-----LTLETFTAINPAPART 105
 Db 805 FAACYSAAVYTLAALGEIFAARAIMSWIGDCAKIIASNDHVPKRWITPLGLPVQVPCRS 864
 QY 106 EVAOKIVRSIKRCKME 122
 Db 865 E--RHILRTSLQVLAQ 879

RESULT 25
 POLG_ZYMVR STANDARD: PRT: 3083 AA.

AC Q089330.
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Genome polypeptide (Contains: N-terminal protein (p1); Helper component polypeptide (EC 3.4.22.45) (HC-Pro); Protein p3; 6 kDa protein 1 (p61); Cytoplasmic inclusion protein (CI); 6 kDa protein 2 (p62); Genome-linked inclusion protein (VPG); Nuclear inclusion protein A (NI-A) (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-Pro); Nuclear inclusion protein B (NI-B) (RNA-directed RNA polymerase) (EC 2.7.7.48); Coat protein (CP)).
 DE Zucchini yellow mosaic virus (strain Reunion Island) (ZYMV).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae; Potyviruses.
 OC NCBI_TaxID=117129;
 RP [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95146958; PubMed=7844540;
 RA Wistler G.C., Purcell D.E., Hiebert E.;
 RT "Characterization of the p1 protein and coding region of the zucchini yellow mosaic virus";
 RL J. Gen. Virol. 76:37-45(1995).
 CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
 CC -1- FUNCTION: CITOPASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT MAY BE INVOLVED IN REPLICATION.
 CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes glutamyl bonds, and activity is further restricted by preferences for the amino acids in p6 - p1, glut(ser or gly) for the enzyme from tobacco etch virus. The natural substrate is the viral polypeptide, but other proteins and oligopeptides containing the appropriate consensus sequence are also cleaved.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate + (RNA)(N).
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes a gly-1-gly bond at its own C-terminus, commonly in the sequence -Tyr-Xaa-Val-Gly-1-gly. In the processing of the polyviral polypeptide.
 CC -1- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
 CC -1- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE POLYPEPTIDE WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT INDIVIDUAL PROTEINS.
 CC -1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
 CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
 CC -1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPEPTIDE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L29569; AAA65558.1; -
 CC HSSP: P27958; IHEI.
 CC MEROPS: C04.003; -.

DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR001730; Peptidase_C4.
 DR InterPro: IPR001456; Peptidase_C6.
 DR InterPro: IPR002540; Poly_P1.
 DR InterPro: IPR001592; Poly_P1.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00271; helicase_C_1.
 DR Pfam: PF00680; RNA_dep_RNA_pol_1.
 DR Pfam: PF00767; Poly_coat_1.
 DR Pfam: PF00851; Peptidase_C6_1.
 DR Pfam: PF00863; Peptidase_C4_1.
 DR Pfam: PF01577; Poly_P1_1.
 DR PRINTS: PR00966; NIAPOVPTASE.
 DR SMART: SM00487; DEXDC_1.
 DR SMART: SM00490; HELIC_1.
 KW Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase; Coat protein; Polypeptide; Covalent protein-RNA linkage; Helicase; ATP-binding.
 KW CHAIN 1 307 N-TERMINAL PROTEIN.
 FT CHAIN 308 769 HELPER COMPONENT PROTEINASE.
 FT CHAIN 770 1115 PROTEIN P3.
 FT CHAIN 1116 1167 6 KDA PROTEIN 1.
 FT CHAIN 1168 1801 CYTOPLASMIC INCLUSION PROTEIN.
 FT CHAIN 1802 1854 6 KDA PROTEIN 2.
 FT CHAIN 1855 2044 GENOME-LINKED PROTEIN.
 FT CHAIN 2045 2287 NUCLEAR INCLUSION PROTEIN A.
 FT CHAIN 2288 2804 NUCLEAR INCLUSION PROTEIN B.
 FT CHAIN 2805 3083 COAT PROTEIN.
 FT NP_BIND 1252 1259 ATP (POTENTIAL).
 SQ SEQUENCE 3083 AA; 351156 MW; 55E51B455C70C537 CRC64;

Query Match 10.2%; Score 69; DB 1; Length 3083;
 Best Local Similarity 23.4%; Pred. No. 2e+02;
 Matches 26; Conservative 18; Mismatches 49; Indels 18; Gaps 4;

QY 18 ASCASTPSPNSKNSANLTSLIKHAVKOTCOTOLGHQWYKIAMKLSSEKAKISTETA 77
 Db 17 ARCATVVGQNGHVN-----IVAGHVAVCKPQKSHSYKHKHSEKSKASEIN--I 66

QY 78 CGCVADKAPK---AVSLTETTAINPAPARTVAVOKIVRSIKRCKMETVN 125
 Db 67 LNSFTDDPEKRFRLTRNEMSKVKKGPGR-----MLRKPRARVLEIRIS 112

RESULT 26
 VGLX_HSVB STANDARD: PRT: 797 AA.
 AC P28968;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-DEC-1992 (Rel. 24, Last annotation update)
 DE Glycoprotein X precursor.
 GN 71.
 OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Alphaherpesvirinae; Varicelloviruses.
 OC NCBI_TaxID=31520;
 RP [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92293566; PubMed=1318606;
 RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
 RT "The DNA sequence of equine herpesvirus-1";
 RL Virology 189:304-316(1992).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----

```

CC EMBL: M8664; AAB02506.1; -
DR PIR: H36802; VGBEX1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 797 GLYCOPROTEIN X.
FT DOMAIN 23 465 SER/THR-RICH.
FT TRANSMEM 766 790 POTENTIAL.
FT CARBOHYD 590 590 N-LINKED (GLCNAC.. ) (POTENTIAL).
SQ SEQUENCE 797 AA; 80342 MW; 50C9ED9211F5E5B2 CRC64;

Query Match 10.1%; Score 68.5; DB 1; Length 797;
Best Local Similarity 25.0%; Pred. No. 46;
Matches 36; Conservative 7; Mismatches 86; Indels 15; Gaps 2;

OY 5 LVTATATLMTAFATLASCAPSPESNPKNSANL-----TSLIKHAVKQCQQLT 53
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 17 IYVIGSTTTTETTTSSSTSGSGOSTSSGTTNSSSPTTSPPTTSSPPSTHTSSST 76
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 54 GHQYKRIAMKLSSESKA---KISETACGCVADKAPAVSLTETLTAIINPAREVAQ 109
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 77 STQSSSTAATSSAPASSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 136
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 110 KIYRHSKRCMLETVAIVPTTT 133
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 137 TTAASSTAETTTATATATPTTT 160
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 27
CAPU_DROME STANDARD; PRT; 1059 AA.
ID CAPU_DROME 024120; Q9VOV8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Capuccino protein.
CN CAPU OR CG3399.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Ovary;
RX MEDLINE=96033799; PubMed=7590229;
RA Emmons S., Phan H., Calley J., Chen W., James B., Manseau L.,
RT "Capuccino, a Drosophila maternal effect gene required for polarity
RT of the egg and embryo, is related to the vertebrate limb deformity
RT locus."
RT Genes Dev. 9:2482-2494(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abilaj J.F., Abmayyal A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernick B.P., Bhandari D., Bolshakov S.,
RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,

```

```

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegyam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., Mobergson D.,
RA Merkulov G., Malshtina N.V., Moberg C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- SIMILARITY: CONTAINS 1 FORMIN HOMOLOG 1 (FH1) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FORMIN HOMOLOG 2 (FH2) DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE FORMIN HOMOLOG FAMILY. CAPPUCCINO
SUBFAMILY.
CC -----
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U34258; AAC46925.1; -
DR EMBL: AE003578; AAF51054.1; -
DR FlyBase; FBgn0000236; capu.
DR InterPro; IPR00104; FH2.
DR InterPro; IPR001265; Formin.
DR Pfam; PF02181; FH2; 1.
DR PRINTS; PR00828; FORMIN.
DR SMART; SM00498; FH2; 1.
KW Developmental protein.
FT DOMAIN 480 560 FH1 (PRO-RICH).
FT DOMAIN 585 1021 FH2.
FT CONFLICT 260 260 S -> C (IN REF. 1).
FT CONFLICT 364 364 S -> T (IN REF. 1).
FT CONFLICT 386 386 T -> S (IN REF. 1).
FT CONFLICT 471 471 E -> K (IN REF. 1).
FT CONFLICT 495 495 H -> P (IN REF. 1).
FT CONFLICT 513 513 MISSING (IN REF. 1).
SQ SEQUENCE 1059 AA; 113863 MW; 009B0E24F61B6E45 CRC64;

Query Match 10.1%; Score 68.5; DB 1; Length 1059;
Best Local Similarity 21.7%; Pred. No. 64;
Matches 33; Conservative 17; Mismatches 51; Indels 51; Gaps 4;

OY 22 STPESNKNSANTTSLIK-----HAV 44
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 250 STWOSNPKSSSGANQELFTLALQFCNNLKYGVCLKOISNEHLDCGSPREMYQWTHTE 309
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 45 KQTCOTOLTHQWIKIAMKLSSE-----SKAKISETACGCVADKAPAVSLTETLTA 97
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 310 QPTTSLPLTFGCKDKVAAWPFSSPTSRALAESLSLADAGVAG-----SLATATA 363
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 98 AINPNAETVAQKIVRHSKPC-MLETVAFT 128
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 364 STASDQKTLQOILKRLNCTTIAEVHAV 395
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 28
MSB2_YEAST STANDARD; PRT; 1306 AA.
ID MSB2_YEAST P32334;

```

```

DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE MSB2 protein (Multicopy suppression of a budding defect 2).
GN MSB2 OR YGR014W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92383951; PubMed=1514328;
RA Bender A., Pringle J.R.
RT "A Ser/Thr-rich multicopy suppressor of a cdc24 bud emergence
RT defect."
RL Yeast 8:315-323(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=5288C;
RX MEDLINE=97435481; PubMed=9290212;
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT chromosome VII."
RL Yeast 13:1077-1090(1997).
CC -1- PPM: O-GLYCOSYLATED IN THE SER/THR-RICH REGIONS (PROBABLE).
CC -1- SIMILARITY: SOME, TO YEAST HKR1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M77354; AAA34798.1; -
DR EMBL: 272799; CAA6997.1; -
DR PIR: S25370; S25370.
DR SGD: S0003246; MSB2.
KM Transmembrane; Glycoprotein; Repeat.
FT DOMAIN 698 816 7 X 17 AA TANDEM REPEATS.
FT REPEAT 698 714 1.
FT REPEAT 715 731 2.
FT REPEAT 732 748 3.
FT REPEAT 749 765 4.
FT REPEAT 766 782 5.
FT REPEAT 783 799 6.
FT REPEAT 800 816 7.
SQ SEQUENCE 1306 AA; 133114 MW; 67D5D984D5CA46D CRC64;

Query Match
Best Local Similarity 10.1%; Score 68.5; DB 1; Length 1306;
Matches 29; Conservative 25; Mismatches 74; Indels 7; Gaps 2;

```

```

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glutamate racemase (EC 5.1.1.3).
GN MURI OR LI1282.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LI1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Winkler P., Manger S., Jallion O., Malarre K.,
RA Weisenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis LI1403."
RL Genome Res. 11:731-753(2001).
CC -1- FUNCTION: PROVIDES THE (R)-GLUTAMIC ACID REQUIRED FOR CELL WALL
CC BIOSYNTHESIS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: L-glutamate -> D-glutamate.
CC -1- PATHWAY: Peptidoglycan biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE ASPARTATE/GLUTAMATE RACEMASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE006360; AAK05380.1; -
DR HSSP; P56868; 1B74.
DR InterPro; IPR001920; Asp/Glu_rac.
DR InterPro; IPR004391; Glu_rac.
DR Pfam; PF01177; Asp_Glu_rac; 1.
DR TIGRfams; TIGR00067; glut_race; 1.
DR PROSITE; PS00923; ASP_GLU_RACEMASE_1; 1.
DR PROSITE; PS00924; ASP_GLU_RACEMASE_2; 1.
KM Peptidoglycan synthesis; Cell wall; Isomerase; Complete proteome.
SQ SEQUENCE 271 AA; 30219 MW; ED575B54565A9D3B CRC64;

Query Match
Best Local Similarity 10.1%; Score 68; DB 1; Length 271;
Matches 30; Conservative 15; Mismatches 42; Indels 18; Gaps 5;

```

```

DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-1999 (Rel. 38, Last annotation update)
DE Major fibrillar subunit precursor.
GN PMFA.
OS Proteus mirabilis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Proteus.
OX NCBI_TaxID=584;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-42.
RX STRAIN=HI4320;
RX MEDLINE=93162827; PubMed=8094384;
RA Bahrani F.K., Cook S., Hull R.A., Massad G., Mobley H.L.T.;
RT "Proteus mirabilis fibrillae: N-terminal amino acid sequence of a
RT major fibrillar subunit and nucleotide sequences of the genes from two

```

```

DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-1999 (Rel. 38, Last annotation update)
DE Major fibrillar subunit precursor.
GN PMFA.
OS Proteus mirabilis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Proteus.
OX NCBI_TaxID=584;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-42.
RX STRAIN=HI4320;
RX MEDLINE=93162827; PubMed=8094384;
RA Bahrani F.K., Cook S., Hull R.A., Massad G., Mobley H.L.T.;
RT "Proteus mirabilis fibrillae: N-terminal amino acid sequence of a
RT major fibrillar subunit and nucleotide sequences of the genes from two

```


FT CARBOHYD 686 686 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHD 879 879 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHD 1092 1092 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHD 1099 1099 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 1169 AA; 122164 MW; 940CDDFB3569C669 CRC64;

Query Match
 Best Local Similarity 25.0%; Score 67.5; DB 1; Length 1169;
 Matches 29; Conservative 20; Mismatches 38; Indels 29; Gaps 6;

OY 13 TAFTIASCASCTPESNPNKSSANLITSLIKHAKVQCOQOLGHOYKTAAMKLSSEKRA 71
 DB 938 TLITVSSCS-----NCSNTVSAVSTATTITNGITTEYTCPLSATETLTYSKL 990
 OY 72 K-----ISETAC--GCVADKAPKPAVSLTETLTAIN-----PNARTEVA 108
 DB 991 ESEKTLITVSSCSGVCSETPASPAVST--ATATVNDVTVYVTSWSPQATINKLA 1044

RESULT 33
 MAPX_DROME

ID MAPX_DROME STANDARD; PRT; 1185 AA.

AC P23226; Q9V9S1;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 205 kDa microtubule-associated protein.
 GN MAP205 OR CGI1483.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RP SEQUENCE FROM N.A.
 RX MEDLINE=9115949; PubMed=1703540.
 RA Irminger-Finger I., Laymon R.A., Goldstein L.S.B.;
 RT "Analysis of the primary sequence and microtubule-binding region of
 J. cell Biol. 111:2563-2572(1990).
 RL [2]

RP SEQUENCE FROM N.A. (ISOFORM B3).
 RC STRAIN=Berkeley;

RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sulten G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazet R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abiri J.F., Agdayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Bottler P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunn B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz G., Ferreira S., Fleischmann W.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegwan C.,
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Jimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Plitman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).

CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE REGULATION OF
 CC MICROTUBULE ASSEMBLY AND INTERACTION.
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH CYTOPLASMIC MICROTUBULES AND
 CC WITH THE MITOTIC SPINDLE.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; C2, B3 AND J5 (SHOWN HERE); ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- MISCELLANEOUS: PHOSPHORYLATION OF VARIOUS SERINE RESIDUES MAY PLAY
 CC A REGULATORY ROLE. THE BASIC DOMAIN CONTAINS NUMEROUS SEQUENCES
 CC THAT MATCH KNOWN CONSENSUS SEQUENCES OF SEVERAL DIFFERENT PROTEIN
 CC KINASES.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X54061; CAA37996.1; -
 CC EMBL; AE003780; AAF57214.1; -
 CC PIR; A36685; A36685.
 CC FLYbase; FBgn002645; Map205.

KW Microtubules; Alternative splicing; Phosphorylation.
 FT DOMAIN 1 784 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 785 1124 ARG/LYS-RICH (BASIC).
 FT BINDING 1125 1185 ASP/GLU-RICH (ACIDIC).
 FT VARSPLIC 745 977 TO MICROTUBULES (POTENTIAL).
 FT VARSPLIC 557 578 MISSING (IN ISOFORM B3 AND ISOFORM C2).
 FT VARSPLIC 650 703 MISSING (IN ISOFORM C2).
 FT VARSPLIC 704 704 D -> N (IN ISOFORM C2).
 SO SEQUENCE 1185 AA; 126669 MW; 47B4222CE03F70 CRC64;

Query Match
 Best Local Similarity 10.0%; Score 67.5; DB 1; Length 1185;
 Matches 34; Conservative 17; Mismatches 51; Indels 33; Gaps 5;

OY 8 TATLTATFTLASCASCTPESN-----PKNSSANLITSLIKHAKVQCOQOLGHOYKTAAM 63
 DB 819 TETLYMKKTTRASSVSYGAKSAAPRSTARLGKISTSIARKTSTTSLTGN-----PRK 873
 OY 64 KLSSEKAKISSETACGCVADKAPKPAVSLTETLTAIN-----PNARTEVAOKI 111
 DB 874 SLSSNWSGV-----KPKTKLSGTRPATPAVSKVYLGAATTINKPTASGASDNY 923
 OY 112 VRHSIKPCMEETVNA 126
 DB 924 TRTTLRP--LVSTNA 936

RESULT 34
 NCA2_XENLA

ID NCA2_XENLA STANDARD; PRT; 1092 AA.

AC P36335;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Neural cell adhesion molecule 2, 180 kDa isoform precursor (N-CAM
 DE 180).
 GN NCAM2.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus.

OX	NCB1_Taxid=8335;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=93273239; PubMed=7684721;	
RA	Tonissen K.F., Kriegl P.A.;	
RT	"Two neural-cell adhesion molecule (NCAM)-encoding genes in Xenopus laevis are expressed during development and in adult tissues.";	
RL	Gene 127:243-247(1993).	
CC	-1- FUNCTION: THIS PROTEIN IS A CELL ADHESION MOLECULE INVOLVED IN NEURON-NEURON ADHESION, NEURITE FASCICULATION, OUTGROWTH OF NEURITES, ETC.	
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.	
CC	-1- ALTERNATIVE PRODUCTS: A number of isoforms are produced by alternative splicing	
CC	-1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.	
CC	-1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.	
CC	-1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch)	
CC	-----	
DR	EMBL; M76710; AAA49910.1; .	
DR	HSP; JN0635; JN0635.	
DR	HSP; P56276; IRLK.	
DR	InterPro: IPR003961; FN_III.	
DR	InterPro: IPR003006; IG_MHC.	
DR	InterPro: IPR003598; IG_C2.	
DR	InterPro: IPR003600; IG_Like.	
DR	Pfam; PF00041; fn3; 2.	
DR	Pfam; PF00047; Ig; 5.	
DR	SMART; SM00060; FN3; 2.	
DR	SMART; SM00410; IG_Like; 1.	
DR	SMART; SM00408; IGC2; 4.	
KW	Cell adhesion; Glycoprotein; Transmembrane; Repeat; Brain; Immunoglobulin domain; Alternative splicing; Signal.	
FT	FT SIGNAL	1 19
FT	FT CHAIN	20 1092
FT	FT DOMAIN	20 705
FT	FT TRANSMEM	706 723
FT	FT DOMAIN	724 1092
FT	FT DOMAIN	34 100
FT	FT DOMAIN	129 193
FT	FT DOMAIN	225 289
FT	FT DOMAIN	316 386
FT	FT DOMAIN	413 480
FT	FT DOMAIN	512 589
FT	FT DOMAIN	618 686
FT	FT DOMAIN	149 153
FT	FT DOMAIN	158 162
FT	FT DISULFID	41 93
FT	FT DISULFID	136 186
FT	FT DISULFID	232 282
FT	FT DISULFID	323 379
FT	FT DISULFID	420 473
FT	FT CARBOHYD	82 82
FT	FT CARBOHYD	219 219
FT	FT CARBOHYD	310 310
FT	FT CARBOHYD	341 341
FT	FT CARBOHYD	417 417
FT	FT CARBOHYD	443 443
FT	FT CARBOHYD	472 472
SO	SEQUENCE	1092 AA; 118082 MW; CD336E0F8B7AD1 CRC64;

```

Oy 7 VTATLMTFAFTLASCASPESNPKSSANLTTSLIKHVAWQOTQTLGHQYKIAAMKLS 66
Db 819 VTTVTNSTLTITETPATKQNSPTSETTILTS-----TAPPTSTADNSTVQSVQ-A 869
Oy 67 SESKAKISETACGCYADDAPEAVSLTELT-----TAINPNAETVAOKIV 112
Db 870 TPSKAEV-PTLASSPPTSPKVPALVLDSDPTNNPSKAVANQAGALNNSAATSAA----- 924
Oy 113 RHSLKPCMLETIVNAFIVPTTT 133
Db 925 -----EPPTALIKPYTT 936

RESULT 35
NXSB_NAJAT STANDARD: PRT: 82 AA.
AC P80958; Q42285;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Short neurotoxin precursor (Cobrotoxin B).
OS Naja atra (Chinese cobra).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Elapidae; Elapinae; Naja.
OX NCBI_TaxID=8656;
RN [1]
RN SEQUENCE FROM N.A., SEQUENCE OF 22-82, AND FUNCTION.
RP TISSUE=Liver, and Venom;
RX MEDLINE=98158338; PubMed=9498573;
RA Chang L.-S., Chou Y.-C., Lin S.-R., Wu B.-N., Lin J., Hong E.,
RA Sun Y.-J., Hsiao C.-D.;
RT "A novel neurotoxin, cobrotoxin b, from Naja naja atra (Taiwan cobra)
RT venom: purification, characterization, and gene organization.";
RL J. Biochem. 122:1252-1259(1997).
RN [2]
RN SEQUENCE FROM N.A., AND REVISION TO 67.
RP TISSUE=Venom gland;
RA Chu R.C., Yang G.C.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: PRODUCES PERIPHERAL PARALYSIS BY BLOCKING NEUROMUSCULAR
CC TRANSMISSION AT THE POSTSYNAPTIC SITE. BINDS TO THE NICOTINIC
CC ACETYLCHOLINE RECEPTOR (BY SIMILARITY).
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its use
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/).license@isb-sib.ch).
CC -----
DR EMBL: Y13399; CAA73829.2; -.
DR EMBL: AF031472; AAB86636.1; -.
DR HSPB; P01427; INOR.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam: PF000087; toxin; 1.
DR Prodom: PD000206; Snake_toxin; 1.
DR PROSITE: PS00272; SNAKE_TOXIN; 1.
KW Venom; Neurotoxin; Postsynaptic neurotoxin; Signal.
FT SIGNAL 1 21
FT CHAIN 22 82 SHORT NEUROTOXIN.
FT DISULFID 24 44 BY SIMILARITY.
FT DISULFID 38 61 BY SIMILARITY.
FT DISULFID 63 74 BY SIMILARITY.
FT DISULFID 75 80 BY SIMILARITY.
SO SEQUENCE 82 AA; 9139 MW; 1EFA21189C08B6E8 CRC64;

Query Match 9.8%; Score 66; DB 1; Length 82;
Best Local Similarity 7.3%; Pred. No. 5.9;
Matches 27; Conservative 12; Mismatches 42; Indels 18; Gaps 4;
2 WKILVYATLMTFAFTLASCASPESNPKSSANLTTSLIKHVAWQOTQTLGHQYKIA 61

```

```

RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Baeil C.R., Ketchum K.A., Lee J.J., Rouning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Unayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Copenhagen G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.,
RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RA thaliana."
RL Nature 402:761-768(1999).
RN 121
RP SEQUENCE FROM N.A.
RC STRAIN:CV, Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGEC).";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC 1.- SIMILARITY: BELONGS TO THE UPF0082 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AC005395; AAC42247.2; -
DR EMBL: AY114086; AAM45134.1; -
DR EMBL: AY056230; AAL07079.1; -
DR InterPro: IPR002876, DUF28.
DR Pfam: PF01709; DUF28; 1.
DR ProDom: PD004323; DUF28; 1.
KW Hypothetical protein.
FT CONFLICT 87
SQ SEQUENCE 331 AA; 36814 MW; 7E3D3D2CABEE1467 CNC64;

Query Match 9.8%; Score 66; DB 1; Length 331;
Best Local Similarity 22.5%; Pred. No. 29;
Matches 32; Conservative 25; Mismatches 67; Indels 18; Gaps 3;

QY 4 ILVATLTATPTATLASCSTPESNPKSSANNTSLIKAAKQTCQQL----- 52
DB 23 ILNSTNRLSLITNTNTSLSSISPHNTTSHFTVAASQSDQDCFRLLQRLKISITPL 82
QY 53 -TGHOYWKIAAKKLSSEK-----AKISFTACGVADAPAVSITTELTAAINPNARTE 106
DB 83 CGMRSSSIACRKGQADSKKAKLYCRICKLEVYNAVKKGPRPVSTNTLAT-ILDAKELD 141
QY 107 VAKQIVRHSLKPCMLMETYNAPL 128
DB 142 VPKDIVERNIKRASEKGGEAFI 163

RESULT 38
C6D2_DROME
ID C6D2_DROME STANDARD; PRT; 512 AA.
AC 09W223;
DE 16-OCT-2001 (Rel. 40, Created)
DE 16-OCT-2001 (Rel. 40, last sequence update)
DE 16-OCT-2001 (Rel. 40, last annotation update)
DE Probable cytochrome P450 6d2 (EC 1.14.-.-) (CYPI4D2).
DE CYP6D2 OR CG4373.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN:Berkeley;
CC VERIFIED BY:

```



```

CC CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd66a.htm"
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC CC at the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR EMBL: X16354; CAA34404.1; -
DR EMBL: J03858; AAA51826.1; -
DR EMBL: D90311; BAA14341.1; -
DR EMBL: D90312; BAA14342.1; -
DR EMBL: D90313; BAA14343.1; -
DR EMBL: M69176; AAA51825.1; -
DR EMBL: M72238; AAA58393.1; -
DR EMBL: M72238; AAA58394.1; -
DR EMBL: AC004785; AAC18434.1; -
DR EMBL: AC004785; AAC18437.1; -
DR EMBL: AC004785; AAC18438.1; -
DR EMBL: AC004785; AAC18439.1; -
DR PIR: A32164; A32164; -
DR PIR: JH0394; JH0394; -
DR PIR: JH0395; JH0395; -
DR PIR: JH0396; JH0396; -
DR GeneW: HGNC:1814; CEACAM1.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR003598; IG_C2.
DR InterPro: IPR003600; IG_Like.
DR Pfam: PF00047; IG; 4.
DR SMART: SM00410; IG_Like; 1.
DR SMART: SM00408; IG_C2; 2.
KW Immunoglobulin domain; Glycoprotein; Transmembrane; Signal;
KW Repeat; Alternative splicing.
FT CHAIN 1 34
FT STGNL 1 34
FT 35 526
FT 35 428
FT 429 452
FT 453 526
FT 35 142
FT 160 222
FT 252 306
FT 341 403
FT 215 299
FT 259 346
FT 346 396
FT 35 35
FT 104 104
FT 111 111
FT 115 115
FT 152 152
FT 182 182
FT 197 197
FT 208 208
FT 224 224
FT 232 232
FT 254 254
FT 274 274
FT 288 288
FT 292 292
FT 302 302
FT 309 309
FT 345 345
FT 351 351
FT 363 363
FT 378 378
FT 405 405
FT 321 321
FT 322 526
FT 321 351
FT 321 351

```

```

FT FT DEAVGQHHPQHKPCOEGGMDLV (IN ISOFORM I).
FT FT MISSING (IN ISOFORM I).
FT FT VARSPLIC 352 526
FT FT VARSPLIC 416 417
FT FT VARSPLIC 418 526
SO SO MISSING (IN ISOFORM G).
SO SO SEQUENCE 526 AA: 57560 MW: 62663 MW: CADIB238D069AF8 CRC64;

Query Match 9.8%; Score 66; DB 1; Length 526;
Best Local Similarity 25.7%; Pred. No. 50;
Matches 29; Conservative 13; Mismatches 45; Indels 26; Gaps 4;

OY 26 SNPKNSANTTSLIKHAVKQCTQTLTGHOY-WKIAAMKLSSEKAKISETAGCGVADK 84
DB 326 AKPOIKASKTYYTGDKDSVNLTCSTNDTGISIRMFKNQSLPSEEMKLSQGN----- 378
OY 85 APEAVSLTELTVAINPNMR-----TEVAQIVRHSLKPCMLETVNFIYP 130
DB 379 -----TTLISNPVAREADAGTYWCEVFNPISKNQSDPIWL-NVNNNALP 420

RESULT 40
GAGT_DROFU
ID GAGT_DROFU STANDARD; PRT; 574 AA.
AC P21331.
DT 01-MAY-1991 (rel. 18; Created)
DT 01-MAY-1991 (rel. 18; Last sequence update)
DT 15-JUN-2002 (rel. 41; Last annotation update)
DE Nucleic-acid-binding protein (mobile element jockey) (ORF1).
GN GAG.
OS Drosophila funebrits (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7221;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1911.1;
RX MEDLINE=91067678; PubMed=1701254;
RA Mizrokhi L.J., Mazo A.M.;
RT "Evidence for horizontal transmission of the mobile element jockey
between distant drosophila species."
RL Proc. Natl. Acad. Sci. U.S.A. 87:9216-9220(1990).
CC -1- FUNCTION: STRONGLY BASIC PROTEIN THAT BINDS DIRECTLY TO
CC RETROVIRAL RNA AND MAY BE INVOLVED IN ITS PACKAGING AND
CC IN THE REVERSE TRANSCRIPTION PROCESS.
CC -1- SIMILARITY: STRONG TO THE EQUIVALENT PROTEIN OF DROSOPHILA
CC MELANOGASTER.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC CC at the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR EMBL: M38437; AAA28648.1; -
DR FlyBase: FBgn0019259; Dfuna\jockey\yag.
DR InterPro: IPR001878; Znf_CCHC.
DR SMART: SM00343; Znf_C2HC; 1.
KW Transposable element; Zinc-finger.
FT DOMAIN 393 447
FT ZN_FING 393 406
FT ZN_FING 412 426
FT ZN_FING 433 447
SO SEQUENCE 574 AA: 62663 MW: 7246722F8A3BF1B7 CRC64;

Query Match 9.8%; Score 66; DB 1; Length 574;
Best Local Similarity 26.4%; Pred. No. 55;
Matches 32; Conservative 10; Mismatches 43; Indels 36; Gaps 5;

OY 7 VTATMTAFTL-----ASCAS-----TPESNPKNSANTTSLI-KH 42
DB 62 VTTSDFSTCTVTTTQSCASTSTTYTVAARVGNICSTTVGN--NSVSDTMAOMAGKA 119

```

Tue Apr 29 12:09:22 2003

us-10-048-196-2.rsp

Page 23

QY 43 AVRCQTQTLTGHQYKIAMKILSSSEKARISSETAGCCVADKAPVLSLTETLTAANIN 102
120 ARKKCCSFQAG-----LSSKKRRVYNACQTIKKVMQMTDLKAKIPKVAATNS 169

QY 103 A 103

Db 170 A 170

Search completed: April 28, 2003, 16:03:52
Job time : 22 secs

This Page Blank (uspto)

Oy 1 MMKILVYATALMTAFITLASASPFES--NRKNSSANLTLSLKHAAVKQTOQTQLTGQYW 58
Db 1 MNKTLSI--LPVALIGGCAAGSGNFTFSLDGTGTGGSGIVKMAAVEQSOCRAELIKRSEW 57
Oy 59 KIAMKISSEBKAISTACGCVAADKPPEAVSLTELTIAINPRAPTEVAOKIVRSHLP 118

QY	DB	Matches	Score	DB	Length	Indels	Gaps
QY	DB	30	13.1%	27.2%	142	15	4
DB	DB	28	Conservative	22	Mismatches	38	Indels
QY	DB	77	ACGCVADAPAVSLTE--LTTA	117	106		
DB	DB	64	TTQMLTYASSNAVSVLNDIMDSKVNP	106			
QY	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	30	NSSANLTTSLIKH-----AVKOTCQ	117			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	77	ACGCVADAPAVSLTE--LTTA	117	106		
DB	DB	64	TTQMLTYASSNAVSVLNDIMDSKVNP	106			
QY	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	30	NSSANLTTSLIKH-----AVKOTCQ	117			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	77	ACGCVADAPAVSLTE--LTTA	117	106		
DB	DB	64	TTQMLTYASSNAVSVLNDIMDSKVNP	106			
QY	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	30	NSSANLTTSLIKH-----AVKOTCQ	117			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	77	ACGCVADAPAVSLTE--LTTA	117	106		
DB	DB	64	TTQMLTYASSNAVSVLNDIMDSKVNP	106			
QY	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	30	NSSANLTTSLIKH-----AVKOTCQ	117			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	77	ACGCVADAPAVSLTE--LTTA	117	106		
DB	DB	64	TTQMLTYASSNAVSVLNDIMDSKVNP	106			
QY	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	30	NSSANLTTSLIKH-----AVKOTCQ	117			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	77	ACGCVADAPAVSLTE--LTTA	117	106		
DB	DB	64	TTQMLTYASSNAVSVLNDIMDSKVNP	106			
QY	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	30	NSSANLTTSLIKH-----AVKOTCQ	117			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	77	ACGCVADAPAVSLTE--LTTA	117	106		
DB	DB	64	TTQMLTYASSNAVSVLNDIMDSKVNP	106			
QY	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	30	NSSANLTTSLIKH-----AVKOTCQ	117			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	77	ACGCVADAPAVSLTE--LTTA	117	106		
DB	DB	64	TTQMLTYASSNAVSVLNDIMDSKVNP	106			
QY	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	30	NSSANLTTSLIKH-----AVKOTCQ	117			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	77	ACGCVADAPAVSLTE--LTTA	117	106		
DB	DB	64	TTQMLTYASSNAVSVLNDIMDSKVNP	106			
QY	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	30	NSSANLTTSLIKH-----AVKOTCQ	117			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	77	ACGCVADAPAVSLTE--LTTA	117	106		
DB	DB	64	TTQMLTYASSNAVSVLNDIMDSKVNP	106			
QY	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	30	NSSANLTTSLIKH-----AVKOTCQ	117			
DB	DB	4	NENDNLTSKQIKFDAMLTPEPTIDKACQ	63			
QY	DB	77	ACGCVADAPAVSLTE--LTTA	117	106		


```
RESULT 5
08TGE1 ID 08TGE1 PRELIMINARY; PRT; 1713 AA.
AC 08TGE1:
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cell wall protein Awalp.
GN AWAL.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-2191416; PubMed-11916725;
RA Shimo H., Sakamoto K., Okuda M., Athi R., Iwashita K., Ito K.;
RT "The AWAL gene is required for the foam-forming phenotype and cell
RT surface hydrophobicity of sake yeast.";
RL Appl. Environ. Microbiol. 68:2018-2025(2002).
DR EMBL; AB071164; BAB85832.1; -.
SQ SEQUENCE 1713 AA; 166883 MW; 9321A296EB586404 CRC64;

Query Match 12.3%; Score 83; DB 3; Length 1713;
Best Local Similarity 26.1%; Pred. No. 18;
Matches 36; Conservative 21; Mismatches 47; Indels 34; Gaps 8;

OY 7 VTATLMTAFTLASCASPESN-----PKNSSANLTSLIKHAVKOTCO--TQLTGH 55
DB 1323 LTVATTTSTVAVSCSETDSDSGNVTYTTTTPCSCS---TTATTTSCDEGCHVTSTG- 1377
OY 56 QYWKIAMKLSSSEKAKISETAC---GC---VADKAPAVSLTELTTAALNPNAETVA 108
DB 1378 ---TVARETVSSKSYTYTVTHCNDNGCNTKTVTSEAPEA-----TTTYSVSP--KITTT 1426
OY 109 QKIVRSLKPCMLETVNA 126
DB 1427 ATVTQCDNCGCKTKVTYS 1444

RESULT 6
09S4J4 ID 09S4J4 PRELIMINARY; PRT; 229 AA.
AC 09S4J4:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Serum opacity factor precursor (Fragment).
GN SOF3018.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-SS1468 M TYPE ST3018;
RA Beall B., Gherardi G., Li Z.;
RT "The relation of Streptococcus pyogenes sof and emm gene sequence
RT types to genetically distinct strain sets.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF139746; AAD42200.1; -.
KW Signal.
FT NON_TER 1 1
FT SIGNAL 12 12
FT CHAIN 13 >229 POTENTIAL.
FT NON_TER 229 229 SERUM OPACITY FACTOR.
SQ SEQUENCE 229 AA; 23738 MW; 7489FCB473CF9236 CRC64;

Query Match 12.2%; Score 82.5; DB 2; Length 229;
Best Local Similarity 25.5%; Pred. No. 2;
Matches 36; Conservative 20; Mismatches 52; Indels 33; Gaps 4;

OY 8 TATLMTAFTLASCASPESNPKNSSANLTSLIKHAVKOTCOTQLTGHQYWKIAMKLSS 67
```

```
DB 34 STTGTSTVTAASESSSSD-----ASVSSGGRQTSASAQSKPOAQTAIV-ASS 83
OY 68 ESKAKISETACGCVADKAPAVSLTELTTAALNP-----NATTEVAOKI----- 111
DB 84 SSSSKANSSSSASADVPKPAVSTTSSATVASFNSGSKANNAETEPQOMMEVEKYTVD 143
OY 112 -----VRHSLKPCMLETVN 125
DB 144 KENSELKVKGDTQPKKGSTVN 164

RESULT 7
09X6G6 ID 09X6G6 PRELIMINARY; PRT; 246 AA.
AC 09X6G6:
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Serum opacity factor precursor (Fragment).
GN SOF8.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-SS634 M TYPE 8;
RA Beall B., Gherardi G., Li Z.;
RT "The relation of Streptococcus pyogenes sof and emm gene sequence
RT types to genetically distinct strain sets.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF138790; AAD31483.1; -.
KW Signal.
FT NON_TER 1 1
FT SIGNAL 13 >246 POTENTIAL.
FT CHAIN 13 >246 SERUM OPACITY FACTOR.
FT NON_TER 246 246
SQ SEQUENCE 246 AA; 25574 MW; 1AEC9EE6C6C56B04 CRC64;

Query Match 12.2%; Score 82.5; DB 2; Length 246;
Best Local Similarity 25.5%; Pred. No. 2.1;
Matches 36; Conservative 20; Mismatches 52; Indels 33; Gaps 4;

OY 8 TATLMTAFTLASCASPESNPKNSSANLTSLIKHAVKOTCOTQLTGHQYWKIAMKLSS 67
DB 34 STTGTSTVTAASESSSD-----ASVSSGGRQTSASAQSKPOAQTAIV-ASS 83
OY 68 ESKAKISETACGCVADKAPAVSLTELTTAALNP-----NATTEVAOKI----- 111
DB 84 SSSSKANSSSSASADVPKPAVSTTSSATVASFNSGSKANNAETEPQOMMEVEKYTVD 143
OY 112 -----VRHSLKPCMLETVN 125
DB 144 KENSELKVKGDTQPKKGSTVN 164

RESULT 8
09X9C0 ID 09X9C0 PRELIMINARY; PRT; 435 AA.
AC 09X9C0:
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Serum opacity factor 2 (Fragment).
GN SOF2.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-A8;
```

RX MEDLINE-99429095; PubMed-10499280;
 RA Kreikemeyer B., Martin D.R., Chhatwal G.S.;
 RT "Sfrit Protein, a fibronectin binding surface protein of group A
 RT streptococci, is a serum opacity factor with high serotype specific
 RT apolipoproteinase activity.";
 RL FEMS Microbiol. Lett. 178:305-311(1999).
 DR EMBL: AJ012315; CAB44274.1;..
 DR InterPro: IPR002335; Myoglobin.
 DR PRINTS: PR002035; VWF_A.
 DR SMART: SM00327; VMA: 1.
 DR PROSITE: PS50234; VMA: 1.
 FT NON_TER 1
 FT 435 435
 SQ SEQUENCE 435 AA; 47916 MW; 3FA70B5DIC541568 CRC64;

Query Match 12.1%; Score 82.5; DB 2; Length 435;
 Best Local Similarity 25.5%; Pred. No. 4.2;
 Matches 36; Conservative 20; Mismatches 52; Indels 33; Gaps 4;

OY 8 TATLMTAFTLASCASTPESNPKNSANLTSLIKHAVKOTCOTOLTGHOYKIAMKISS 67
 DB 3 STTGTSTVTAASEASSESD-----ASVSSGGRQTSESAQASKOPAQTA--ASS 52
 OY 68 ESKAKISFTACGVADKAPKAVSLTETLTA--NARTEVAOKI----- 111
 DB 53 SSSSKAKISSSSADYAPKAVSTSSSATVAPSPNSGSKNEKAMNETEQOMVEKITYD 112
 OY 112 -----VRHSLKPCMLETVN 125
 DB 113 KENSELKVKDCTQPKRGTVN 133

RESULT 9
 O9CELO PRELIMINARY; PRT; 179 AA.

AC 09CELO: 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Unknown protein.
 OS YSUF OR LI1827.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Lactococcus.
 OX NCBI_Taxid=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-114403.
 RX MEDLINE-21235186; PubMed-11337471;
 RA Bolotin A., Winchester P., Mauger S., Jallion O., Malarme K.,
 RT Weissenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis ssp. lactis 11403.";
 RL Genome Res. 11:731-753(2001).
 DR EMBL: AE006413; AAK05925.1; .
 KW Complete proteome.
 SQ SEQUENCE 179 AA; 19145 MW; 05A7ADA2AD4AD65 CRC64;

Query Match 12.1%; Score 82; DB 16; Length 179;
 Best Local Similarity 23.1%; Pred. No. 1.6;
 Matches 33; Conservative 21; Mismatches 55; Indels 34; Gaps 4;

OY 1 MKLIVYATATMTAFTLASCASTPESNPKNSANLTSLIKHAVKOTCOTOLTGHOYKI 60
 DB 1 MKKLITTTTALALSLGACSKKSDASHKSSSTSSSPSSGTS----- 45
 OY 61 AAKLSESKAKISFTACGVADKAPKAVSLTETLTA--NARTEVAOKI 111
 DB 46 SSSSKAKISSSSK--ETPGNIDSSFOKAVDNTQTVATMKETKYDTSIDTATAEAPQTI 102
 OY 112 VRHSLKPCMLETVN 134

DB 103 V-----YTYTYRNQITPTATK 118

RESULT 10

O8TDH7 PRELIMINARY; PRT; 1029 AA.
 AC 08TDH7: 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Intestinal membrane mucin MUC17 (Fragment).
 GN MUC17.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE-21845432; PubMed-11855812;
 RA Gum J.R. Jr., Crawley S.C., Hicks J.W., Szymkowski D.E., Kim Y.S.;
 RT "MUC17, a novel membrane-tethered mucin.";
 RL Biochem. Biophys. Res. Commun. 291:466-473(2002).
 DR EMBL: AF430017; AAL89737.1; .
 FT NON_TER 1
 FT 1029 1029
 SQ SEQUENCE 1029 AA; 108320 MW; 5E1AD19EABF948E3 CRC64;

Query Match 12.1%; Score 82; DB 4; Length 1029;
 Best Local Similarity 28.8%; Pred. No. 13;
 Matches 38; Conservative 15; Mismatches 61; Indels 18; Gaps 6;

OY 8 TATLMTAFTLASCASTPESNPKNSANLTSLIKHAVKOTCOTOLTGHOYKIAMKISS 67
 DB 534 TAAPLTYYTM-----STAPSTPRTSROGCTTS-----ASTISATSTPRTSVTTPVTP 583
 OY 66 SSKAKISFTACGVADKAPKAVSLTETLTA--INPAREVAOKIYRHSKPCML 121
 DB 584 SSES-SRPSSTTSHPTTPPPHASTPTASSTVNPVAVTMTTR--TKPSTRITTSF 641
 OY 122 ETVAFTVPTTT 133
 DB 642 PYVTITTAAPTNT 653

RESULT 11

O9LH95 PRELIMINARY; PRT; 608 AA.
 AC 09LH95: 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Genomic DNA, chromosome 3, BAC clone: T19N8.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-COLUMBIA;
 RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
 RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN-COLUMBIA;
 RX MEDLINE-20363099; PubMed-10907853;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
 RT Sequence features of the regions of 4,251,695 bp covered by ninety pl.
 RT TAC and BAC clones.";
 RL DNA Res. 7:217-221(2000).
 DR EMBL: AF002057; BAB03177.1; .
 DR InterPro: IPR002965; P-rich_extensn.
 DR PRINTS: PR01217; PRICHEXTNSN.

```

SQ      SEQUENCE      608 AA;   62766 MW;   6C6368AF4BD2A3B9 CRC64;

Query Match          11.8%; Score 80; DB 10; Length 608;
Best Local Similarity 26.8%; Pred. No. 11;
Matches 37; Conservative 25; Mismatches 50; Indels 26; Gaps 6

OY      18 ASASIPTEENPKNSSANLTTSLIKHAIVKQT-----COTQLTGHQYWKI----- 60
        ::::| | | | : | : | : | | | | | | | | : | : | : | |
DB      420 SAASQRKEKNSKSSSSSSSTTSVKEVETOSSEVNFSISNLEKKYGNSELKVFEEKLT 479
        ::::| | | | : | : | : | | | | | | | | : | : | : | |
OY      61 ---AAAKLTSESASKISEFGACGVADKAPEAVSLFELTTAAINPNARTEVA--QKTVRHS 115
        ::::| | | | : | : | : | | | | | | | | : | : | : | |
DB      480 SMSASAKKLST-SNAKELVTGMRSAASKIAEAMVEVS-SRFKSSEETKTSMASCQOEVMQS 537
        ::::| | | | : | : | : | | | | | | | | : | : | : | |
OY      I16 LKPCMLETVNAFLVPTTT 133
        || | : | : | | |
DB      538 LKE--LDQINSQIVSGKT 553
        || | : | : | | |

RESULT 12

ID      Q12955      PRELIMINARY;      PRT; 4377 AA.
AC      Q12955;
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      Ankyrin G.
       GN
       ANK-3.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
RX      NCB_I_TaxId=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE-BRAIN STEM;
RX      MEDLINE=95138209; PubMed=7836469;
RT      Kordell E., Lambert S., Bennett V.;
       "Ankyrin: A new ankyrin gene with neural-specific isoforms localized
       at the axonal initial segment and node of Ranvier." ;
RL      J. Biol. Chem. 270:2352-2359(1995).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE-BRAIN STEM;
RA      Submitter S.S.;
RL      Submitted (Aug-1994) to the EMBL/GenBank/DDBJ databases.
DR      EMBL: U13616; AAA64834.1; -.
DR      HSP: P55273; IB18.
DR      InterPro: IPR002110; ANK.
DR      InterPro: IPR000488; Death.
DR      InterPro: IPR000906; ZUS.
DR      Pfam: PF00023; ank; 24.
DR      Pfam: PF00531; death; 1.
DR      Pfam: PF00791; ZUS; 1.
DR      PRINTS: PRO1415; ANKYRN.
DR      SMART: SM00248; ANK; 21.
DR      SMART: SM00005; DEATH; 1.
DR      SMART: SM00218; ZUS; 1.
DR      PROSITE: PSS0088; ANK_REPEAT; 21.
DR      PROSITE: PSS0297; ANK_REP_REGION; 1.
DR      PROSITE: PSS0017; DEATH_DOMAIN; 1.
DR      ANK repeat; Repeat.
SO      SEQUENCE 4377 AA; 480395 MW; F42379E55768B684 CRC64;

Query Match          11.8%; Score 79.5; DB 4; Length 4377;
Best Local Similarity 26.7%; Pred. No. 1.2e+02;
Matches 28; Conservative 18; Mismatches 52; Indels 7; Gaps 4;

OY      14 AFLTASCASTPESNP-KNSSANLTTSLIKHAIVKQCOTQLTGHWYKIAAMKLSSESKAK 72
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      3900 ALTTSCCVDAKSRIPVNTPRD---NLII--AVRKACATOKOG-QPERGKAKQLPSKLPVK 3953
        : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      73 ISFTAGCGVADKAPEAVSLFELTTAAINPNARTEVAQKIYRHSLK 117
        : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db      3954 VR8CVTTTATTTTATTTTATTTTCTVAVKRSQKCEVKSIE 3998

RESULT 13
Q25402
ID 025402; PRELIMINARY; PRT; 354 AA.
AC
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, last annotation update)
DE Microfilarial sheath protein SHP3.
GN SHP3.
OS Litomosoides sigmodontis.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Litomosoides.
OX NCBI_TaxID=42156;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=95140057; PubMed=7838185;
RA Hintz M., Hitzmann J., Hobom G., Linder D., Lottspeich F.,
RA Scholt H.H., Conraths F.J., Zahner H., Stilm S.;
RT "Litomosoides carinii microfilarial sheaths: partial amino acid
RT sequences of several major polypeptide constituents.";
RL Mol. Biochem. Parasitol. 67:69-78(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Hitzmann J., Hintz M., Shrestha T.R., Conraths F.J., Stilm S.,
RA Zahner H., Hobom G.;
RT "Ser/Thr-rich repeats in two major microfilarial surface proteins of
RT the filariae Brugia malayi and Litomosoides sigmodontis.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; U54556; AAB82017.2;
SQ SEQUENCE 354 AA; 35850 MW; 8CF7BC5773087688 CRC64;

Query Match 11.7%; Score 79; DB 5; Length 354;
Best Local Similarity 24.1%; Pred. No. 7.3;
Matches 32; Conservative 20; Mismatches 75; Indels 6; Gaps 4;

QY 2 MKILVYATATLMTAFTLASCASPTESNPKRSSANLTLSLIKHAVKOTQQLGHQYKIA 61
Db 1 MKEVLDVIVAFELL--CV-TPRSQASSSTSTTPAK--TTSPTTAKTKSTTKTYY 55
QY 62 AMKLSESKAKISETACGCGVADKAPAEVSLPELTTLAIPNPARTEVAOKIVRHSILKPCML 121
Db 56 TVKSTSTTKTSTSTSTSTTKPTTKTSTTKTSTTKTSTTKTSTTKTSTTKTSTTKT 114
QY 122 ETVNAEIVPTTTR 134
Db 115 STSTSTTKPTTSK 127

RESULT 14
Q9L178
ID Q9L178 PRELIMINARY; PRT; 327 AA.
AC Q9L178;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
DE Possible binding-protein-dependent transport 11proteins.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.

```

```
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A., Denapalte D., Eichner A., Cullum J.,
RT "A set of ordered cosmid and a detailed genetic and physical map for
RL the 8 Mb Streptomyces coelicolor A3(2) chromosome.",
RN Mo. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rahlbowski E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RL coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL: AL158057; CAB76300.1;
DR InterPro: IPR002491; Peripla_BP.
DR Pfam: PF01497; Peripla_BP_2; 1.
KW Lipoprotein.
SQ
SEQUENCE 327 AA; 34094 MW; 49998D116B0CFE75 CRC64;

Query Match 11.5%; Score 78; DB 16; Length 327;
Best Local Similarity 25.7%; Pred. No. 8.4;
Matches 43; Conservative 19; Mismatches 65; Indels 40; Gaps 7;

QY 4 ILVYATATMTAFLLASCST---PESNPKNSSA-----N 34
DB 5 LTRAAATTAALTLACGTPEAPADKAEKASEITLTKDGKTEVKLDGPATKVVATEWN 64
QY 35 LITSLIHAVKQTCQTLTGHOYWKIAAMKLSSEK--AKISPTACGCVDKAPE-AVSL 91
DB 65 VVSLVSLGVDPVGVADVKGKGTMD-SAVPLKNPKDIDIGRGEPSMTPTVASLADPLIVAT 123
QY 92 TELTTAINNARTEVNAQIVRHSLKPC-----MLEFTVNAFTVPTT 133
DB 124 TDLAPAAV--KQLREVAAPVIEKRSADGTGQIDRLMLENVDLIAEATGT 168

RESULT 15
O8TBM5
ID O8TBM5 PRELIMINARY; PRT; 579 AA.
AC O8TBM5;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 61.1 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC022270; AAH22270.1;
KW Hypothetical protein.
SQ
SEQUENCE 579 AA; 61053 MW; 7011F548BE82DA40 CRC64;

Query Match 11.5%; Score 78; DB 4; Length 579;
Best Local Similarity 25.5%; Pred. No. 16;
Matches 25; Conservative 24; Mismatches 45; Indels 4; Gaps 1;
```

```
QY 18 ASCASTPESNPKNSSANLITSLIKHAVKQTCQTLTGHOYWKIAAMKLSSEKAKISPTA 77
DB 323 SSEASVSSSVAKNSSSGTSLTPKSSSTNTSLTSTKSTQVAAASLAKSSSSQS--- 379
QY 78 CGCVADKAPAEVSLTELTAAINPNA RTEVAQIVRHS 115
DB 380 -GSLVSKSTSLASVSQSLAKSSQSTSTSLPSTQS 416

RESULT 16
O9NXV6
ID O9NXV6 PRELIMINARY; PRT; 580 AA.
AC O9NXV6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CDNA FLJ20036 f1s, clone COL00219.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Ohtani R., Ota T., Suzuki Y., Oiyashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK000043; BAA90902.1;
DR InterPro: IPR001159; DS_RBD.
DR PROSITE: PS50137; DS_RBD; 1.
SQ
SEQUENCE 580 AA; 61124 MW; 5CFB5DFEE50A475 CRC64;

Query Match 11.5%; Score 78; DB 4; Length 580;
Best Local Similarity 25.5%; Pred. No. 16;
Matches 25; Conservative 24; Mismatches 45; Indels 4; Gaps 1;

QY 18 ASCASTPESNPKNSSANLITSLIKHAVKQTCQTLTGHOYWKIAAMKLSSEKAKISPTA 77
DB 324 SSEASVSSSVAKNSSSGTSLTPKSSSTNTSLTSTKSTQVAAASLAKSSSSQS--- 380
QY 78 CGCVADKAPAEVSLTELTAAINPNA RTEVAQIVRHS 115
DB 381 -GSLVSKSTSLASVSQSLAKSSQSTSTSLPSTQS 417

RESULT 17
O9NYHO
ID O9NYHO PRELIMINARY; PRT; 631 AA.
AC O9NYHO;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Putative serine-rich protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Duncan E.L., Kaul S.C.;
RT "Homo sapiens mRNA for putative serine-rich protein.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF246705; AAF68967.1;
DR InterPro: IPR001159; DS_RBD.
DR PROSITE: PS50137; DS_RBD; 1.
KW NON-TER
SQ
SEQUENCE 631 AA; 66301 MW; D4023AA68084B56 CRC64;

Query Match 11.5%; Score 78; DB 4; Length 631;
Best Local Similarity 25.5%; Pred. No. 18;
Matches 25; Conservative 24; Mismatches 45; Indels 4; Gaps 1;
```


RC STRAIN-972H-;
 RA Wood V., Warren T., Harris D., Barrell B.G., Rajadream M.A.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AU591302; CAC38347.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 3971 AA; 394861 MW; 5408B39FD97E735D CRC64;

Query Match
 Best Local Similarity 11.5%; Score 77.5; DB 3; Length 3971;
 Matches 35; Conservative 12; Mismatches 46; Indels 35; Gaps 6;

QY 6 YVATLTMTAFTLASCASPESNPKNSANLTTSLIKHAKQTCOTQLTGHQYWKIAMKL 65
 DB 272 YIT-TLQTTTAAQ--TISENTYSTGPNLTSTNSPQISST-----I 311
 QY 66 SESEKAKISERACGVADKAPKAVSLTELTTAALNPARTENAKIVRSIKRCMLETYN 125
 DB 312 SSSS-----FIVESPSVALSTISSTTTT---MASTPAANTLISRSKP--TDTTN 356
 QY 126 AFTVPTT 133
 DB 357 SISFANTT 364

RESULT 22
 025437 PRELIMINARY; PRT; 415 AA.
 ID 025437;
 AC 025437;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBlrel. 08, Last annotation update)
 DE Nanos homolog.
 GN MD NOS.
 OS Musca domestica (House fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Muscidae; Muscidae; Musca.
 RX NCBI_TaxID=7370;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=91347366; PubMed=1908748;
 RT Wang C., Lehmann R.;
 RL "Nanos is the localized posterior determinant in Drosophila.";
 RN Cell 66:637-647(1991).
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95324413; PubMed=7601003;
 RT Curtis D., Apfeld J., Lehmann R.;
 RL "Nanos is an evolutionarily conserved organizer of anterior-posterior polarity.";
 RN Development 121:1899-1910(1995).
 DR EMBL: U24696; AAA87461.1; -;
 SQ SEQUENCE 415 AA; 44935 MW; 02747E23570648BE CRC64;

Query Match
 Best Local Similarity 11.4%; Score 77; DB 5; Length 415;
 Matches 19; Conservative 12; Mismatches 44; Indels 0; Gaps 0;

QY 8 TATLTMTAFTLASCASPESNPKNSANLTTSLIKHAKQTCOTQLTGHQYWKIAMKLSS 67
 DB 227 TPTSLTAANQOONGTSATSPKSTINPTVALTPPOOLQOHINMSENFHWKILPAHMOO 286
 QY 68 ESKAKISERACGVA 82
 DB 287 HSHAAVTAAAAAAA 301

RESULT 23
 09G0H8 PRELIMINARY; PRT; 540 AA.
 AC 09G0H8;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE GP7.
 OS Rosephage S101.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
 OC T7-like viruses.
 NCBI_TaxID=136084;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rohwer F.L., Segall A.M., Steward G., Seguritan V., Breitbart M.,
 RA Wolven F., Azam F.;
 RT "The complete genomic sequence of the marine phage Rosephage S101 shares homology with normative phages.";
 RL J. Virol. 76:408-418(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Rohwer F.L., Segall A.M., Steward G., Seguritan V., Breitbart M.,
 RA Wolven F., Azam F.;
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF189021; AAG02590.1; -;
 DR InterPro: IPR000104; Antifreeze_1.
 DR InterPro: IPR004089; Chitaxis_transd.
 DR PRINTS: PR00308; ANTIFREEZE.
 DR PRINTS: PR00833; POAALLERGEN.
 DR PRINTS: PR01574; TUBBYPROTEIN.
 SQ SEQUENCE 540 AA; 51936 MW; 450329FFA7A9B09D CRC64;

Query Match
 Best Local Similarity 11.4%; Score 77; DB 9; Length 540;
 Matches 28; Conservative 8; Mismatches 37; Indels 20; Gaps 2;

QY 13 TATLTASCASPESNPKNSANLTTSLIKHAKQTCOTQLTGHQYWKIAMKLSESKAK 72
 DB 184 TATQASAAVSATNAATASNAATASATNAASQ-----AAATSETNAA 228
 QY 73 ISE-----TACGVADKAPKAVSLTELTTAALN 100
 DB 229 ASETVTTSATNAATSEANATASATATTQATN 261

RESULT 24
 008294 PRELIMINARY; PRT; 967 AA.
 ID 008294;
 AC 008294;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Chromosome XV reading frame ORF YOL155C.
 GN YOL155C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 RX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gallion L., Dujon B.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Arino J., Casamayor A., Gano F.J., Gancedo C., Lafuente M.J.,
 RA Aldea M., Casas C., Herrero E.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MIPS;
 DR Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z74897; CAA99177.1; -;
 DR SGD: S0005515; YOL155C.
 DR PRINTS: PR01608; BACINVASING.
 SQ SEQUENCE 967 AA; 94704 MW; 7BFC01EA243A561E CRC64;

Query Match
 Best Local Similarity 11.4%; Score 77; DB 3; Length 967;
 Matches 24.3%; Pred. No. 38;

Matches 34; Conservative 21; Mismatches 57; Indels 28; Gaps 7;

QY 7 VITFL--MTAFTLASCAT-----PESNPKNSANLTTSLIKHAVKOTCOTQ-LT 53
 DB 709 ITTVCSSSTTATITSCDEGCHVTSTGTVAETVSSKSYTVTVTHCDNNGCNTKTVT 768
 QY 54 GHYWKIAAMKLSSESKAKISFAC--GC----VADKAEVSLTETLTAIPNARTE 106
 DB 769 SECPESATTTSPKSYTVTVTHCDNNGCNTKTVTSEAPEA-----TTTVSP--KTY 820
 QY 107 VAOKIVRHSILKPCMLETVNA 126
 DB 821 TTATVTQCDNCGSTKTVTS 840

RESULT 25

005164 PRELIMINARY; PRT; 1001 AA.

AC 005164; PRELIMINARY; PRT; 1001 AA.
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE AOB567, AOF1001, AOE110, AOE264 and AOE130 genes.
 GN YOL155C OR AOP1001.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FY1679;
 RA Gamo F.J., Latuente M.J., Casamayor A., Aldea M., Casas C., Arlo J.,
 RA Herrero E., Gancedo C.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X89715; CA61860.1; -
 DR SGD: S0005515; YOL155C.
 SQ SEQUENCE 1001 AA; 97673 MW; F11BC4522DCFB4F CRC64;

Query Match 11.4%; Score 77; DB 3; Length 1001;

Best Local Similarity 24.3%; Pred. No. 39;

Matches 34; Conservative 21; Mismatches 57; Indels 28; Gaps 7;

QY 7 VITFL--MTAFTLASCAT-----PESNPKNSANLTTSLIKHAVKOTCOTQ-LT 53
 DB 743 ITTVCSSSTTATITSCDEGCHVTSTGTVAETVSSKSYTVTVTHCDNNGCNTKTVT 802
 QY 54 GHYWKIAAMKLSSESKAKISFAC--GC----VADKAEVSLTETLTAIPNARTE 106
 DB 803 SECPESATTTSPKSYTVTVTHCDNNGCNTKTVTSEAPEA-----TTTVSP--KTY 854
 QY 107 VAOKIVRHSILKPCMLETVNA 126
 DB 855 TTATVTQCDNCGSTKTVTS 874

RESULT 26

010441 PRELIMINARY; PRT; 262 AA.

AC 010441; PRELIMINARY; PRT; 262 AA.
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Putative 30 kDa protein.
 GN L2 COMP. ORF.
 OS Mouse adenovirus type 1 (MAV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OX NCBI_TaxID=10530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Meisner J.D., Hirsch G.N., Larue E.A., Fulcher R.A., Spindler R.R.;
 RT "Completion of the DNA sequence of mouse adenovirus type 1: sequence
 of E2B, L1, and L2 (18-51 map units).";
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: U95843; AAB53756.1; -
 SQ SEQUENCE 262 AA; 30266 MW; 78A1242A3EFCF5E6 CRC64;

Query Match 11.3%; Score 76.5; DB 12; Length 262;

Best Local Similarity 26.5%; Pred. No. 9.2;

Matches 22; Conservative 14; Mismatches 26; Indels 21; Gaps 4;

QY 19 SCASP-ESNP-KNSANLTTSLIKHAVKOTC-----QTQLTGHWYKIAAMK 64
 DB 46 STATPRPTNPSRMSVSLNTWVAHPYGRNCSTRGRCRMPSCNTIMVSHSYKDHVYC 105
 QY 65 LSSE-----SKAKISFACGC 80
 DB 106 VNSHCSARHPPASASYNLTSQCG 128

RESULT 27

052259 PRELIMINARY; PRT; 414 AA.

AC 052259; PRELIMINARY; PRT; 414 AA.
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Trag protein (Fragment).
 GN TRAG.
 OS Escherichia coli.
 OC Plasmid PED208.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-86223783; PubMed-3011738;
 RA Finlay B.B., Paranchych W.;
 RT "Nucleotide sequence of the surface exclusion genes tras and trar from
 the incF0 lac Plasmid PED208.";
 RL J. Bacteriol. 166:713-721(1986).
 DR EMBL: M13465; AAA88373.1; -
 KW Plasmid.
 FT NON_TER
 SQ SEQUENCE 414 AA; 44120 MW; 3447BFDA601125C CRC64;

Query Match 11.3%; Score 76.5; DB 2; Length 414;

Best Local Similarity 29.4%; Pred. No. 16;

Matches 32; Conservative 16; Mismatches 46; Indels 15; Gaps 5;

QY 26 SNPKNSANL-----TSLIKHAVKOTCOTQLTGHWYKIA--AMKLSSESKAKISFETA 77
 DB 192 SHTDNASIASLANOLSYFSSLSKASQASQYNDVATVRSHEY SOLASAEENNSASINONYSOEF 251
 QY 78 CGGVADKAP-EAVSLTETLTAIPNARTE---VAOKIVRHSILKPCMLE 122
 DB 252 VGIVTSKRPEGA---DOLLSDAASPEVRAERDRLAQAFVEDRMKPOLLO 297

RESULT 28

09N8N9 PRELIMINARY; PRT; 513 AA.

AC 09N8N9; PRELIMINARY; PRT; 513 AA.
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Probable axoneme central apparatus proteinh.
 GN CHR1.235.
 OS Trypanosoma brucei.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5691;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TREU927;
 RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,
 RA Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,
 RA Gerard C., Rajandream M.A., Barrell B.G.;

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL359782; CAB95504.1; -
 DR InterPro: IPR000225; Armadillo.
 DR Pfam: PF00514; Armadillo_seg. 8.
 DR SMART: SM00185; ARM; 6.
 SO SEQUENCE 513 AA; 56126 MW; 40BB4546A1B390C3 CRC64;

Query Match 11.3%; Score 76.5; DB 5; Length 513;
 Best Local Similarity 23.3%; Pred. No. 20;
 Matches 30; Conservative 12; Mismatches 44; Indels 43; Gaps 4;

OY 17 LASCSTPSNPKNSANLTSLIKH-----AVKQTCTQLTGHQYWKIAAMKLSSES 69
 DB 172 LVLCVQEPFLSKRTAASTLADIANKLPDLAQAVVDQDAVTHLA-----PLIGSN 221
 OY 70 KAKISETACGVADKAPAVSLTEL-----TTAANPNART----- 105
 DB 222 DGLKRYCCGCAQIAKNSVELAEVGEIPIRIFSLIKSDSETVRKNASTREVAKH 281
 OY 106 --EVAQKIV 112
 DB 282 TPELAQLIV 290

RESULT 29

OY082 PRELIMINARY; PRT; 515 AA.
 AC O97082;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Hypothetical 55.9 kDa protein.
 GN T39A15.70 OR A74G27580.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC NCBI_TaxID=3702; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;

RA Bevan M., Van Der Schueren J., Chuang Y.-J., Voet M., Robben J.,
 RA Volckaert G., Hohelsel J., Mewes H.W., Mayer K.F.X., Schueller C.,
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Robben J., Gijmompres B., Volckaert G., Mewes H.W., Lemcke K.,
 RA Mayer K.F.X.,
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]

RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL035602; CAB8270.1; -
 DR EMBL: AL161571; CAB81408.1; -
 DR InterPro: IPR001107; Band 7.
 DR InterPro: IPR001972; Stomatin.
 DR Pfam: PF01145; Band 7; 1.
 DR PRINTS: PR00721; STOMATIN.
 DR SMART: SM00244; PHB; 1.
 KW Hypothetical protein.
 SO SEQUENCE 515 AA; 55923 MW; 420FE3DB5B7AFDBA CRC64;

Query Match 11.2%; Score 76; DB 10; Length 515;
 Best Local Similarity 29.3%; Pred. No. 23;
 Matches 27; Conservative 15; Mismatches 44; Indels 6; Gaps 3;

OY 16 TLASCSTP-ESNPKNSANLTSLI--KHAVKOTCQQLTGHQYWKIAAMKLSSESRAK 72
 DB 411 TMGCAKPKRESDLVESVSTENAVESKNAATETDAITLQEKKEESIEETKKEGETFED 470

OY 73 ISETACGVADKAPAVSLTELTTAANPNAR 104
 DB 471 SSEAT---KAEPPEAVKAEKTSSETEPPAQ 499

RESULT 30

OY082 PRELIMINARY; PRT; 189 AA.
 AC O82RB1;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Glycoprotein/polysaccharide metabolism.
 GN YBAY OR STM0465.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC NCBI_TaxID=602;

RP SEQUENCE FROM N.A.
 RC STRAIN-LT2 / SGC1412 / ATCC 700720;
 RX MEDLINE-21534948; PubMed-11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Portolillo S., All J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.,
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2."
 RL Nature 413:852-856(2001).
 DR EMBL: AE008717; AL19420.1; -
 KW Complete proteome.
 SO SEQUENCE 189 AA; 19477 MW; E2708C011CB1E1B9 CRC64;

Query Match 11.2%; Score 75.5; DB 16; Length 189;
 Best Local Similarity 26.7%; Pred. No. 7.9;
 Matches 39; Conservative 21; Mismatches 47; Indels 39; Gaps 9;

OY 2 MKLIYVATLTMTAFTLASCA-----STPSNPKNSANLTSLIKHNAVQTCOTQLTGH 55
 DB 1 MKLVHIVSGLAIVALSACADKSDIQTAPNP-NMSITANQSHIQ-----QPNVSG- 51
 OY 56 QYW---KIA-----AMKSSSEKAK-----ISETACGVADKAPAVSLTELTTAAN 100
 DB 52 TWVIRQVALPDAVLVTLSLSDASLADAPSKVLSQKAVRTGKQADPSFVL-PNPSDIQ 110
 OY 101 PNARTEVAQKI-----VRHSIKP 118
 DB 111 PNARILSAITVDNKLVTITDSVKP 136

RESULT 31

OY082 PRELIMINARY; PRT; 189 AA.
 AC O82B02;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical lipoprotein STY0509.
 GN STY0509.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC NCBI_TaxID=601;

RP SEQUENCE FROM N.A.
 RC STRAIN-C718;
 RX MEDLINE-21534947; PubMed-11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher S., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaiha M.,
 RA Baker S., Basham K., Brooks K., Chillingworth T., Conerton P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,


```

DR HSP6: P00736; IAO.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001507; Endoglin/CD105.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF_15.
DR SMART: SM00179; EGF_Ca_12.
DR SMART: SM00001; EGF_like_16.
DR SMART: SM00241; Zp_1.
DR PROSITE: PS00010; ASX_HYDROXYL; 9.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 22.
DR PROSITE: PS01187; EGF_Ca_13.
DR PROSITE: PS00135; TRYPSIN_SER; UNKNOWN_1.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 3507 AA; 36515 MW; 154F0B68787AD9DF CRC64;

Query Match 11.0%; Score 74.5; DB 5; Length 3507;
Best Local Similarity 21.6%; Pred. No. 3, 1e+02;
Matches 30; Conservative 30; Mismatches 50; Indels 29; Gaps 4;

QY 13 TAFTLASCATPESPENPKNSANLTLTLKHAVKOTQOTGLGQYKIAAKISS----- 67
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2449 TSNVLSSSTTPRESSKSPVSSSTEGT-----SVTSTFSKVEPSTSTSYLED 2498

QY 68 -----ESKAKISETACGACVADKAEVASLTETLTAAIN---PNARTEVAOKIVRH 114
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2499 LRTTTSPLLEETTTTASSETSEPLETDSLVSVRIHLLTTSSENVKRESSTTTSSSKRP 2558
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 115 SLKPKMEYNAVIVPTT 133
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2559 SQEPACILT-STVAVPTSS 2576

RESULT 36
Q82068
ID 082068 PRELIMINARY: PRT: 298 AA.
AC 082068;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Glycoprotein (Fragment).
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SUBGROUP A.
RX MEDLINE=94335057; PubMed=8057427;
RA Garcia O., Martin M., Dopazo J., Arbliza J., Fabrasile S., Russi J.,
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
RA Melero J.A.;
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
RT A): cocirculating lineages and correlation of genetic and antigenic
RT changes in the G glycoprotein."
RL J. Virol. 68:5448-5459(1994).
DR EMBL: Z33429; CA83872.1; -.
DR InterPro: IPR000925; Glycoprot_G.
DR InterPro: IPR003880; Pentamete_attach.
DR Pfam: PF00802; Glycoprotein_G_1.
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 298 AA; 32753 MW; 7D080307897A772B CRC64;

Query Match 10.9%; Score 74; DB 12; Length 298;
Best Local Similarity 24.3%; Pred. No. 19;
Matches 27; Conservative 20; Mismatches 50; Indels 14; Gaps 3;

1 MKKLTVAATLMAFTLASCATPESPENPKNSANLTLTLKHA---VKKOTQOTGLGQYKIAAKISS-----

```


DB 503 EBEKEDLRVQLKRNHPSPLPGSKTSKRPIKIVSLISQGDPTAGPCAPSGSAPEAAGK 562
 QY 90 -----SLTELTTAINPNARTEVA---OKIVRHSILKPCMLETVNARIVPT 131
 DB 563 PITMTLIGASAGAKELTGLTTAKSSSSSEGGVSASPVPSVSSSTAPSAHTLIQSRIVAT 622
 QY 132 T 132
 DB 623 S 623

RESULT 40

Q9HFS1 PRELIMINARY; PRT; 250 AA.
 AC Q9HFS1;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE Glycosyl-phosphatidylinositol protein.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Pedros B., Martinez J., Casanova M.;
 RT "Identification and characterization of glycosyl-phosphatidylinositol
 RL proteins (GPI-proteins) of Candida albicans.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF307520; AAC29538.1; -
 SO SEQUENCE 250 AA; 25163 MM; C6227CD31F011DD8 CRC64;

Query Match 10.8%; Score 73; DB 3; Length 250;
 Best Local Similarity 25.8%; Pred. No. 19;

Matches 33; Conservative 19; Mismatches 54; Indels 22; Gaps 4;

QY 7 VTATLMTAFITLASCSTPESNPKNSANLTTSLIKHAVKOTQOTLGHQYWKIA---A 62
 DB 126 VSDALAKAANAASATTTSTATKSAAEELAT-----SDTTIVASTSHES-KVAETSVA 178
 QY 63 MKLSSESKAKISETACGCVADKPEAVSLTELTTAINPNARTEVAQKIVRHSILKPCMLE 122
 DB 179 QQTASTEKSSAETSRAKETSKEAESKAE-----TSVAOSSSSANVAYSVAE 227
 QY 123 TVNAFTVP 130
 DB 228 TANAGNMP 235

Search completed: April 28, 2003, 16:04:27
 Job time : 38 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 16:04:58 ; Search time 35 Seconds
(without alignments)
510.160 Million cell updates/sec

Title: US-10-048-196-2

Perfect score: 134
Sequence: 1 MKKILYVATIMTAFTLASC.....SLKPCMLETVAFIYPTTTR 134

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : A_Geneseq_101002.*

```

1: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	134	100.0	134	22	ABAB0646
2	7	5.2	32	21	AA89752
3	7	5.2	32	21	ABAB01160
4	7	5.2	32	22	AAU13706
5	7	5.2	32	22	AA878153
6	7	5.2	40	22	ABAB02611
7	7	5.2	62	22	AA882470
8	7	5.2	77	21	AA601641
9	7	5.2	87	23	ABP33369
10	7	5.2	131	21	AA601640

11	7	5.2	147	21	AA655444	Arabidopsis thalia
12	7	5.2	154	22	AB623244	Novel human diago
13	7	5.2	157	21	AA655443	Arabidopsis thalia
14	7	5.2	170	23	AAU12129	Arabidopsis RPW8.1
15	7	5.2	170	23	AAU12130	Arabidopsis RPW8.1
16	7	5.2	205	20	AAV37276	Protein which is s
17	7	5.2	215	22	ABAB0422	Drosophila melanog
18	7	5.2	345	22	AAU35475	Haemophilus influe
19	7	5.2	435	21	AA624041	Arabidopsis thalia
20	7	5.2	440	21	AA624040	Arabidopsis thalia
21	7	5.2	466	21	AA636664	Arabidopsis thalia
22	7	5.2	478	21	AA636663	Arabidopsis thalia
23	7	5.2	478	21	AA636611	Arabidopsis thalia
24	7	5.2	618	21	AA636610	Arabidopsis thalia
25	7	5.2	618	23	ABAB3742	Arabidopsis thalia
26	7	5.2	618	23	ABAB4809	Herbicidially activ
27	7	5.2	992	19	AA689740	Listeria monocytog
28	7	5.2	992	19	AA689742	SAPAP1 protein. H
29	7	5.2	1848	22	ABAB6896	SAPAP2 protein. H
30	7	5.2	1848	22	ABAB6896	Drosophila melanog
31	7	5.2	11	17	AA694612	Protein kinase bin
32	7	5.2	11	23	AAU12145	Arabidopsis RPW8.2
33	7	5.2	26	20	AAV36488	Fragment of human
34	6	4.5	29	23	ABG40230	Peptide #4847 enco
35	6	4.5	34	21	AA632011	Human peptide enco
36	6	4.5	39	22	AA680045	Human secreted pro
37	6	4.5	42	22	ABAB35095	Chemokine peptide
38	6	4.5	42	22	ABAB20514	Peptide #2601 enco
39	6	4.5	42	22	AA658217	Protein #2513 enco
40	6	4.5	42	22	AA658217	Human brain expres
41	6	4.5	42	22	AA658217	Human bone marrow
42	6	4.5	42	22	AA658217	Peptide #2544 enco
43	6	4.5	42	22	ABAB37829	Peptide #2637 enco
44	6	4.5	42	22	AA658217	Human peptide enco
45	6	4.5	42	22	AA658217	Alnus glutinosa al
46	6	4.5	42	22	AA658217	Propionibacterium
47	6	4.5	42	22	AA658217	Propionibacterium
48	6	4.5	42	22	AA658217	Human ORX protein
49	6	4.5	42	22	AA658217	Staphylococcus aur
50	6	4.5	42	22	AA658217	Human ORX protein
51	6	4.5	42	22	AA658217	Neisserial conserv
52	6	4.5	42	22	AA658217	Arabidopsis thalia
53	6	4.5	42	22	AA658217	Human colon cancer
54	6	4.5	42	22	AA658217	Human immune/haema
55	6	4.5	42	22	AA658217	Fragment of human
56	6	4.5	42	22	AA658217	Human ORFX protein
57	6	4.5	42	22	AA658217	Human foetal prote
58	6	4.5	42	22	AA658217	Propionibacterium
59	6	4.5	42	22	AA658217	Enterococcus faeca
60	6	4.5	42	22	AA658217	Peptide #9665 enco
61	6	4.5	42	22	AA658217	Peptide #11701 enco
62	6	4.5	42	22	AA658217	Protein #7717 enco
63	6	4.5	42	22	AA658217	Protein #9072 enco
64	6	4.5	42	22	AA658217	Human brain expres
65	6	4.5	42	22	AA658217	Human brain expres
66	6	4.5	42	22	AA658217	Human bone marrow
67	6	4.5	42	22	AA658217	Human bone marrow
68	6	4.5	42	22	AA658217	Peptide #7148 enco
69	6	4.5	42	22	AA658217	Peptide #8258 enco
70	6	4.5	42	22	AA658217	Peptide #10003 enco
71	6	4.5	42	22	AA658217	Peptide #12185 enco
72	6	4.5	42	22	AA658217	Human peptide enco
73	6	4.5	42	22	AA658217	Human peptide enco
74	6	4.5	42	22	AA658217	Human peptide enco
75	6	4.5	42	22	AA658217	Human digestive sy
76	6	4.5	42	22	AA658217	Human secreted pro
77	6	4.5	42	22	AA658217	Human secreted pro
78	6	4.5	42	22	AA658217	Enterococcus faeca
79	6	4.5	42	22	AA658217	Human immune/haema
80	6	4.5	42	22	AA658217	Arabidopsis thalia
81	6	4.5	42	22	AA658217	Arabidopsis thalia
82	6	4.5	42	22	AA658217	Propionibacterium
83	6	4.5	42	22	AA658217	Human ORFX protein

84	6	4.5	93	22	AAU29725	Novel human secret	157	6	4.5	171	21	AAI5558	Arabidopsis thaliana
85	6	4.5	95	22	ABG08511	Novel human diagno	158	6	4.5	176	22	ABG26650	Novel human diagno
86	6	4.5	95	22	ABP06908	Human OREF protein	159	6	4.5	176	22	AAH40867	Human polypeptide
87	6	4.5	96	21	AAI13390	Arabidopsis thaliana	160	6	4.5	178	22	AAI74058	Human colon cancer
88	6	4.5	98	21	AAI35631	Arabidopsis thaliana	161	6	4.5	181	22	AAI63653	Propionibacterium
89	6	4.5	99	23	ABP34896	Human helicase-11k	162	6	4.5	181	22	AAU64584	Propionibacterium
90	6	4.5	99	23	ABP01565	Human OREF protein	163	6	4.5	182	22	AAI31895	Amilo acid sequenc
91	6	4.5	99	23	ABP09719	Human OREF protein	164	6	4.5	183	22	ABBI5288	Human nervous syst
92	6	4.5	100	21	AAI35630	Arabidopsis thaliana	165	6	4.5	184	22	ABBI5288	Human nervous syst
93	6	4.5	100	22	AAU21491	Human novel foetal	166	6	4.5	184	22	ABBI5288	Human nervous syst
94	6	4.5	101	23	ABP32700	Human OREF1673 prot	167	6	4.5	191	21	AAI3910	Arabidopsis thaliana
95	6	4.5	104	20	AAI34659	Chlamydia pneumonia	167	6	4.5	191	21	AAI3910	Arabidopsis thaliana
96	6	4.5	105	22	AAI90511	Human immune/haema	168	6	4.5	192	23	ABP10588	Human OREF protein
97	6	4.5	115	22	AAI90511	Human immune/haema	169	6	4.5	193	23	AAI90588	Human OREF protein
98	6	4.5	115	22	AAI90511	Human immune/haema	170	6	4.5	195	20	AAI90588	Human OREF protein
99	6	4.5	116	21	AAI90511	Human immune/haema	171	6	4.5	195	20	AAI90588	Human OREF protein
100	6	4.5	119	23	ABP06674	Eucalyptus grandis	172	6	4.5	195	20	AAI90588	Human OREF protein
101	6	4.5	122	22	ABG09337	Human OREF protein	173	6	4.5	195	20	AAI90588	Human OREF protein
102	6	4.5	122	22	ABP10527	Human OREF protein	174	6	4.5	195	20	AAI90588	Human OREF protein
103	6	4.5	125	22	AAU27996	Human contig polyp	175	6	4.5	199	22	AAI31893	Amilo acid sequenc
104	6	4.5	130	13	AAI22370	Novel human diagno	176	6	4.5	199	22	AAI31893	Amilo acid sequenc
105	6	4.5	131	13	AAI22370	Novel human diagno	177	6	4.5	201	22	AAI78597	Human lung tumour
106	6	4.5	131	20	AAI89318	Stimac339 tat gene	178	6	4.5	201	22	AAI78597	Human lung tumour
107	6	4.5	131	20	AAI89318	Stimac339 tat gene	179	6	4.5	201	22	AAI78597	Human lung tumour
108	6	4.5	133	23	AAI16542	UDP-glucuronosyltr	180	6	4.5	202	18	AAI25548	Human lung tumour
109	6	4.5	133	22	AAI17646	Human nervous syst	181	6	4.5	203	22	AAI25548	Human lung tumour
110	6	4.5	134	22	AAU40585	Propionibacterium	182	6	4.5	203	22	AAI25548	Human lung tumour
111	6	4.5	136	22	AAI70692	S cerevisiae apopt	183	6	4.5	204	23	AAI25548	Human lung tumour
112	6	4.5	140	21	AAI25419	Pinus radiata cell	184	6	4.5	204	23	AAI25548	Human lung tumour
113	6	4.5	140	21	AAI25419	Pinus radiata cell	185	6	4.5	204	23	AAI25548	Human lung tumour
114	6	4.5	143	22	ABG18420	Human polypeptide	186	6	4.5	205	22	AAI25548	Human lung tumour
115	6	4.5	144	11	AAI06431	Novel human diagno	187	6	4.5	205	22	AAI25548	Human lung tumour
116	6	4.5	144	11	AAI06431	Novel human diagno	188	6	4.5	210	22	AAI25548	Human lung tumour
117	6	4.5	144	22	ABG18516	N-terminal of SPL-	189	6	4.5	211	22	AAI25548	Human lung tumour

376	6	4.5	429	21	AAI75809	Neisseria gonorrhoe	449	6	4.5	506	23	ABBS4798	Lactococcus lactis
377	6	4.5	429	21	AAI75883	Neisseria strain f	450	6	4.5	507	22	ABG21063	Novel human diagno
378	6	4.5	429	22	ABAB6713	Putative P. abyssi	451	6	4.5	512	22	ABG17742	Novel human diagno
379	6	4.5	431	19	AAW56289	Babesia microti BM	452	6	4.5	518	23	AAU11020	Novel human diagno
380	6	4.5	431	20	AAW24355	Babesia microti an	453	6	4.5	518	23	AAU11020	Novel human diagno
381	6	4.5	431	21	ABG30204	B. microti BMN1-9	454	6	4.5	525	18	AAW24230	Novel human diagno
382	6	4.5	431	23	ABBS8950	Streptococcus poly	455	6	4.5	525	23	ABBS7353	Human melanoma ass
383	6	4.5	431	23	ABBS8950	Streptococcus poly	456	6	4.5	526	9	AAI80986	Mouse ischemic co
384	6	4.5	433	22	ABP29337	Nicotiana tabacum	457	6	4.5	526	13	AAI80986	Sequence of fusion
385	6	4.5	433	22	AAI72079	Nicotiana tabacum	458	6	4.5	526	14	AAI80986	Sequence of fusion
386	6	4.5	436	14	AAI66695	E. coli thymidine	459	6	4.5	526	14	AAI80986	Sequence of fusion
387	6	4.5	436	22	ABG21024	Novel human diagno	460	6	4.5	526	14	AAI80986	Sequence of fusion
388	6	4.5	437	22	AAU14030	Peptide sequence f	461	6	4.5	526	14	AAI80986	Sequence of fusion
389	6	4.5	440	14	AAI66695	E. coli thymidine	462	6	4.5	526	14	AAI80986	Sequence of fusion
390	6	4.5	440	22	ABBS60461	Drosophila melanog	463	6	4.5	526	14	AAI80986	Sequence of fusion
391	6	4.5	440	22	AAU34882	E. coli cellular p	464	6	4.5	526	14	AAI80986	Sequence of fusion
392	6	4.5	440	22	ABBS67583	Botulinum neurotox	465	6	4.5	526	14	AAI80986	Sequence of fusion
393	6	4.5	440	22	ABBS67583	Botulinum neurotox	466	6	4.5	526	14	AAI80986	Sequence of fusion
394	6	4.5	441	23	ABBS69080	S. epidermidis ope	467	6	4.5	526	14	AAI80986	Sequence of fusion
395	6	4.5	445	22	ABBS69070	Neisseria gonorrhoe	468	6	4.5	526	14	AAI80986	Sequence of fusion
396	6	4.5	445	22	ABBS69070	Neisseria gonorrhoe	469	6	4.5	526	14	AAI80986	Sequence of fusion
397	6	4.5	451	21	AAI74914	Neisseria meningit	470	6	4.5	526	14	AAI80986	Sequence of fusion
398	6	4.5	451	23	ABBS6916	Neisseria meningit	471	6	4.5	526	14	AAI80986	Sequence of fusion
399	6	4.5	454	23	AAI645534	Neisseria meningit	472	6	4.5	526	14	AAI80986	Sequence of fusion
400	6	4.5	454	23	AAI645534	Neisseria meningit	473	6	4.5	526	14	AAI80986	Sequence of fusion
401	6	4.5	454	23	AAI645534	Neisseria meningit	474	6	4.5	526	14	AAI80986	Sequence of fusion
402	6	4.5	455	21	AAI645534	Neisseria meningit	475	6	4.5	526	14	AAI80986	Sequence of fusion
403	6	4.5	457	22	ABG21061	Neisseria meningit	476	6	4.5	526	14	AAI80986	Sequence of fusion
404	6	4.5	467	22	AAI645534	Neisseria meningit	477	6	4.5	526	14	AAI80986	Sequence of fusion
405	6	4.5	467	22	AAI645534	Neisseria meningit	478	6	4.5	526	14	AAI80986	Sequence of fusion
406	6	4.5	471	21	AAI645534	Neisseria meningit	479	6	4.5	526	14	AAI80986	Sequence of fusion
407	6	4.5	471	21	AAI645534	Neisseria meningit	480	6	4.5	526	14	AAI80986	Sequence of fusion
408	6	4.5	471	21	AAI645534	Neisseria meningit	481	6	4.5	526	14	AAI80986	Sequence of fusion
409	6	4.5	473	21	AAI645534	Neisseria meningit	482	6	4.5	526	14	AAI80986	Sequence of fusion
410	6	4.5	474	21	AAI645534	Neisseria meningit	483	6	4.5	526	14	AAI80986	Sequence of fusion
411	6	4.5	475	18	AAW24228	Human melanoma ass	484	6	4.5	526	14	AAI80986	Sequence of fusion
412	6	4.5	475	21	AAI645534	Neisseria meningit	485	6	4.5	526	14	AAI80986	Sequence of fusion
413	6	4.5	475	21	AAI645534	Neisseria meningit	486	6	4.5	526	14	AAI80986	Sequence of fusion
414	6	4.5	479	22	ABG21060	Neisseria meningit	487	6	4.5	526	14	AAI80986	Sequence of fusion
415	6	4.5	485	21	AAI645534	Neisseria meningit	488	6	4.5	526	14	AAI80986	Sequence of fusion
416	6	4.5	485	21	AAI645534	Neisseria meningit	489	6	4.5	526	14	AAI80986	Sequence of fusion
417	6	4.5	485	22	AAI645534	Neisseria meningit	490	6	4.5	526	14	AAI80986	Sequence of fusion
418	6	4.5	486	21	AAI645534	Neisseria meningit	491	6	4.5	526	14	AAI80986	Sequence of fusion
419	6	4.5	486	21	AAI645534	Neisseria meningit	492	6	4.5	526	14	AAI80986	Sequence of fusion
420	6	4.5	487	17	AAI645534	Neisseria meningit	493	6	4.5	526	14	AAI80986	Sequence of fusion
421	6	4.5	487	17	AAI645534	Neisseria meningit	494	6	4.5	526	14	AAI80986	Sequence of fusion
422	6	4.5	488	21	AAI645534	Neisseria meningit	495	6	4.5	526	14	AAI80986	Sequence of fusion
423	6	4.5	488	21	AAI645534	Neisseria meningit	496	6	4.5	526	14	AAI80986	Sequence of fusion
424	6	4.5	488	21	AAI645534	Neisseria meningit	497	6	4.5	526	14	AAI80986	Sequence of fusion
425	6	4.5	488	21	AAI645534	Neisseria meningit	498	6	4.5	526	14	AAI80986	Sequence of fusion
426	6	4.5	488	21	AAI645534	Neisseria meningit	499	6	4.5	526	14	AAI80986	Sequence of fusion
427	6	4.5	488	21	AAI645534	Neisseria meningit	500	6	4.5	526	14	AAI80986	Sequence of fusion
428	6	4.5	488	21	AAI645534	Neisseria meningit	501	6	4.5	526	14	AAI80986	Sequence of fusion
429	6	4.5	488	22	AAI645534	Neisseria meningit	502	6	4.5	526	14	AAI80986	Sequence of fusion
430	6	4.5	488	22	AAI645534	Neisseria meningit	503	6	4.5	526	14	AAI80986	Sequence of fusion
431	6	4.5	488	22	AAI645534	Neisseria meningit	504	6	4.5	526	14	AAI80986	Sequence of fusion
432	6	4.5	488	22	AAI645534	Neisseria meningit	505	6	4.5	526	14	AAI80986	Sequence of fusion
433	6	4.5	488	22	AAI645534	Neisseria meningit	506	6	4.5	526	14	AAI80986	Sequence of fusion
434	6	4.5	488	22	AAI645534	Neisseria meningit	507	6	4.5	526	14	AAI80986	Sequence of fusion
435	6	4.5	488	22	AAI645534	Neisseria meningit	508	6	4.5	526	14	AAI80986	Sequence of fusion
436	6	4.5	488	22	AAI645534	Neisseria meningit	509	6	4.5	526	14	AAI80986	Sequence of fusion
437	6	4.5	488	22	AAI645534	Neisseria meningit	510	6	4.5	526	14	AAI80986	Sequence of fusion
438	6	4.5	488	22	AAI645534	Neisseria meningit	511	6	4.5	526	14	AAI80986	Sequence of fusion
439	6	4.5	488	22	AAI645534	Neisseria meningit	512	6	4.5	526	14	AAI80986	Sequence of fusion
440	6	4.5	488	22	AAI645534	Neisseria meningit	513	6	4.5	526	14	AAI80986	Sequence of fusion
441	6	4.5	488	22	AAI645534	Neisseria meningit	514	6	4.5	526	14	AAI80986	Sequence of fusion
442	6	4.5	488	22	AAI645534	Neisseria meningit	515	6	4.5	526	14	AAI80986	Sequence of fusion
443	6	4.5	488	22	AAI645534	Neisseria meningit	516	6	4.5	526	14	AAI80986	Sequence of fusion
444	6	4.5	488	22	AAI645534	Neisseria meningit	517	6	4.5	526	14	AAI80986	Sequence of fusion
445	6	4.5	488	22	AAI645534	Neisseria meningit	518	6	4.5	526	14	AAI80986	Sequence of fusion
446	6	4.5	488	22	AAI645534	Neisseria meningit	519	6	4.5	526	14	AAI80986	Sequence of fusion
447	6	4.5	488	22	AAI645534	Neisseria meningit	520	6	4.5	526	14	AAI80986	Sequence of fusion
448	6	4.5	488	22	AAI645534	Neisseria meningit	521	6	4.5	526	14	AAI80986	Sequence of fusion

522	6	4.5	667	21	AAV54268	Amino acid sequenc	595	6	4.5	987	22	ABB61740	Drosophila melanog
523	6	4.5	671	22	AAB93307	Human protein sequ	596	6	4.5	1003	21	AAB69279	HIV-1 non-subtype
524	6	4.5	674	22	AAB95675	Human protein sequ	597	6	4.5	1013	19	AAW61539	Human cardiac/brai
525	6	4.5	679	21	AAB16685	Bacteriophage Dp-1	598	6	4.5	1013	19	AAW40224	Human tolloid-like
526	6	4.5	681	22	ABG14781	Novel human diago	599	6	4.5	1015	23	ABB93834	Herbicidally activ
527	6	4.5	685	21	AAV93366	Amino acid sequenc	600	6	4.5	1023	23	AAU82954	Human homologue of
528	6	4.5	685	22	AAB19876	Glybon leukemia vi	601	6	4.5	1036	23	ABR46866	Listeria monocytog
529	6	4.5	690	22	ABB60360	Drosophila melanog	602	6	4.5	1040	21	AAAG30935	Arabidopsis thalia
530	6	4.5	691	22	AAU04891	Micromonospora eve	603	6	4.5	1052	21	AAAG30934	Arabidopsis thalia
531	6	4.5	692	22	ABB61668	Drosophila melanog	604	6	4.5	1060	16	AAW40123	Arabidopsis thalia
532	6	4.5	696	22	ABB66411	Drosophila melanog	605	6	4.5	1064	22	AAW41117	IL8-R type 1-GBP 1
533	6	4.5	707	22	AAW23715	Human EST encoded	606	6	4.5	1069	21	AAV93339	A murine interphot
534	6	4.5	707	22	AAW23744	Human EST encoded	607	6	4.5	1070	22	AAW00894	Human bone marrow
535	6	4.5	710	22	ABB59347	Drosophila melanog	608	6	4.5	1094	21	AAW23858	Haemophilus influe
536	6	4.5	718	22	ABG20363	Novel human diago	609	6	4.5	1110	21	AAW08517	Amino acid sequenc
537	6	4.5	727	17	AAR88390	Human neurotranami	610	6	4.5	1120	22	AAU15076	Protein encoded by
538	6	4.5	727	23	AAE21800	Human H1PHUM 00000	611	6	4.5	1125	19	AAW52288	Rattus norvegicus
539	6	4.5	729	22	ABG25904	Novel human diago	612	6	4.5	1139	19	AAW37779	Novel human protei
540	6	4.5	737	15	AAW35494	Novel human diago	613	6	4.5	1162	22	ABB62516	Drosophila melanog
541	6	4.5	737	17	AAW06084	TATA-binding prote	614	6	4.5	1167	23	ABP27982	Streptococcus poly
542	6	4.5	737	18	AAW25019	Human TATA-binding	615	6	4.5	1167	23	ABP27982	Human TBC-1 protei
543	6	4.5	738	20	AAV01070	TATA-binding prote	616	6	4.5	1168	22	AAW36497	Arabidopsis thalia
544	6	4.5	739	23	ABW54873	Human I(3)mbt prot	617	6	4.5	1168	22	AAW36497	Arabidopsis thalia
545	6	4.5	748	22	ABB62227	Lactococcus lactis	618	6	4.5	1181	21	AAW30933	Arabidopsis thalia
546	6	4.5	749	22	ABG18513	Drosophila melanog	619	6	4.5	1187	22	ABB63669	Arabidopsis thalia
547	6	4.5	755	23	ABB93440	Novel human diago	620	6	4.5	1193	23	ABB90860	Drosophila melanog
548	6	4.5	753	22	AAW40484	Herbicidally activ	621	6	4.5	1215	22	AAE02355	Herbicidally activ
549	6	4.5	771	22	AAU69462	Human polypeptide	622	6	4.5	1225	19	AAW52289	Pestivirus NS5AB f
550	6	4.5	772	20	AAV01069	Human purified sec	623	6	4.5	1235	22	AAW59132	Homo sapiens cdo t
551	6	4.5	773	22	ABB65667	Human I(3)mbt prot	624	6	4.5	1240	22	ABB59132	Drosophila melanog
552	6	4.5	773	22	ABB65667	Drosophila melanog	625	6	4.5	1242	19	AAW52287	Rattus norvegicus
553	6	4.5	774	19	AAW35390	Flavobacterium odo	626	6	4.5	1266	22	AAW78463	Keltus norvegicus
554	6	4.5	779	22	ABB57786	Drosophila melanog	627	6	4.5	1276	22	ABW11806	Novel human diago
555	6	4.5	793	22	AAU27571	Neisseria meningit	628	6	4.5	1277	22	ABB71133	Novel human diago
556	6	4.5	793	22	AAE10016	N. meningitidis st	629	6	4.5	1289	22	AAW79447	Drosophila melanog
557	6	4.5	795	21	AAV96995	Glycosylphosphatid	630	6	4.5	1335	22	ABB71593	Human protein SEQ
558	6	4.5	795	23	ABB92756	Herbicidally activ	631	6	4.5	1342	22	ABB64632	Drosophila melanog
559	6	4.5	803	22	AAW31992	Amino acid sequenc	632	6	4.5	1372	19	AAW56473	Drosophila melanog
560	6	4.5	804	21	AAW42128	Human OREF ORF1892	633	6	4.5	1449	23	ABB48252	Protein with Rho p
561	6	4.5	814	21	AAW29830	Arabidopsis thalia	634	6	4.5	1498	22	ABB63547	Drosophila melanog
562	6	4.5	815	23	ABP35606	Fungal ZBC protein	635	6	4.5	1539	22	AAW78908	Listeria monocytog
563	6	4.5	823	21	AAW29829	Arabidopsis thalia	636	6	4.5	1596	22	AAW28354	Human protein SEQ
564	6	4.5	823	21	AAW29829	Arabidopsis thalia	637	6	4.5	1617	21	AAW52174	Novel human secret
565	6	4.5	832	22	AAU27569	Neisseria meningit	638	6	4.5	1633	21	AAW52173	Arabidopsis thalia
566	6	4.5	832	22	AAE10014	N. meningitidis st	639	6	4.5	1637	22	AAU28166	Novel human secret
567	6	4.5	833	20	AAW94058	Murine MSH5 (mMSH5	640	6	4.5	1657	22	ABG17957	Novel human diago
568	6	4.5	840	21	AAV96994	Glycosylphosphatid	641	6	4.5	1683	21	AAW52172	Arabidopsis thalia
569	6	4.5	843	23	ABB08522	Amino acid sequenc	642	6	4.5	1728	22	AAW39234	Human polypeptide
570	6	4.5	853	23	ABB60193	Drosophila melanog	643	6	4.5	1766	22	AAW79892	Human polypeptide
571	6	4.5	856	22	AAU14094	Peptide sequence f	644	6	4.5	1766	22	AAW79892	Human polypeptide
572	6	4.5	858	22	AAU27574	Neisseria meningit	645	6	4.5	1790	22	ABG11807	Novel human diago
573	6	4.5	858	22	AAU27574	N. meningitidis st	646	6	4.5	1815	22	ABB68268	Drosophila melanog
574	6	4.5	880	21	AAW29828	Arabidopsis thalia	647	6	4.5	1931	21	AAW23258	Human polypeptide
575	6	4.5	880	21	AAW29828	Putative P. abyssi	648	6	4.5	1981	21	ABB61657	Human polypeptide
576	6	4.5	883	22	ABG14730	Novel human diago	649	6	4.5	1981	22	AAW40014	Human polypeptide
577	6	4.5	884	22	ABW1494	Drosophila melanog	650	6	4.5	1982	22	AAW34207	Staphylococcus aur
578	6	4.5	887	21	AAW44304	Tomato beta galact	651	6	4.5	2040	22	ABW34207	Staphylococcus aur
579	6	4.5	891	22	AAW25806	Human protein sequ	652	6	4.5	2040	22	ABW34207	Staphylococcus aur
580	6	4.5	897	22	AAU27572	Neisseria meningit	653	6	4.5	2183	21	AAW37018	Drosophila melanog
581	6	4.5	897	22	AAU27572	N. meningitidis st	654	6	4.5	2304	22	ABB60219	Staphylococcus aur
582	6	4.5	901	22	ABB58224	Drosophila melanog	655	6	4.5	2437	22	AAU34358	Murine factor V SE
583	6	4.5	901	22	ABB58224	Drosophila melanog	656	6	4.5	2748	22	ABB58843	Drosophila melanog
584	6	4.5	903	22	ABB62171	Drosophila melanog	657	6	4.5	3158	22	AAU37018	Staphylococcus aur
585	6	4.5	918	7	AAW60723	Sequence of extrac	658	6	4.5	3418	18	AAW19211	Human breast cance
586	6	4.5	927	8	AAW70768	Epstein-Barr virus	659	6	4.5	3418	20	AAW23287	Human breast and o
587	6	4.5	930	21	AAW18212	Plasmodium falcipe	660	6	4.5	3418	20	AAW04355	Human BRCA2 (om12)
588	6	4.5	937	22	AAW38698	Human polypeptide	661	6	4.5	3418	20	AAW04356	Human BRCA2 (om13)
589	6	4.5	942	21	AAW35536	Eucalyptus grandis	662	6	4.5	3418	20	AAW04357	Human BRCA2 (om14)
590	6	4.5	949	23	ABP35671	Fungal ZBC protein	663	6	4.5	3418	20	AAW04358	Human BRCA2 (om15)
591	6	4.5	954	22	AAU35645	Haemophilus influe	664	6	4.5	3418	20	AAW04354	Human BRCA2 (om11)
592	6	4.5	964	22	AAW39331	Human polypeptide	665	6	4.5	3423	22	ABG23417	BRCA2 protein sequ
593	6	4.5	974	22	AAW93175	Human protein sequ	666	6	4.5	3853	22	AAW31167	Novel human diago
594	6	4.5	975	19	AAW69801	Amino acid sequenc	667	6	4.5	3896	21	AAW53617	Amino acid sequenc
			981	21	AAW43195	Human OREF ORF2959							

668	6	4.5	3988	21	AAV53615	NADL protein encod	741	5	3.7	10	22	ABP23280	HIV A11 motif pol
669	6	4.5	3988	21	AAV53616	Amino acid sequenc	742	5	3.7	10	22	ACG88208	Saccharomyces cere
670	6	4.5	4623	22	ABB71106	Drosophila melanog	743	5	3.7	10	23	AAE20209	Human cartilage g1
671	6	4.5	4741	22	AAU31612	Novel human secret	744	5	3.7	10	23	ABB51388	Human 34P3D7 HLA A
672	6	4.5	4743	22	AAU31619	Novel human secret	745	5	3.7	10	23	ABB51388	Human 34P3D7 HLA A
673	6	4.5	5024	22	AAAG82935	S. epidermidis ope	746	5	3.7	10	23	ABB51695	Human 34P3D7 HLA A
674	6	4.5	5292	22	AAU31617	Novel human secret	747	5	3.7	10	23	ABB51703	Human 34P3D7 HLA A
675	6	4.5	6281	22	AAU37403	Staphylococcus aur	748	5	3.7	10	23	ABB51820	Human 34P3D7 HLA B
676	6	4.5	9222	22	ABG21064	Novel human diagno	749	5	3.7	11	15	AAE55537	Human 34P3D7 HLA B
677	6	4.5	10182	23	ABP38314	Staphylococcus epi	750	5	3.7	11	15	AAE55537	Human 34P3D7 HLA B
678	5	3.7	7	17	AAV01953	Mutated fragment o	751	5	3.7	11	17	AAW06895	Anti-CD18 Fab epit
679	5	3.7	7	22	AAU03540	Synthetic heptapep	752	5	3.7	11	18	AAW27332	Salvage receptor b
680	5	3.7	7	23	ABP48646	zinc finger protei	753	5	3.7	11	18	AAW30194	Salvage receptor b
681	5	3.7	7	23	ABP48673	zinc finger protei	754	5	3.7	11	18	AAW34507	Salvage receptor b
682	5	3.7	7	23	ABP48685	zinc finger protei	755	5	3.7	11	18	AAW24059	Salvage receptor b
683	5	3.7	7	23	ABP48688	zinc finger protei	756	5	3.7	11	19	AAW70627	Salvage receptor b
684	5	3.7	7	23	ABP49125	zinc finger protei	757	5	3.7	11	19	AAW62020	Salvage receptor b
685	5	3.7	7	23	ABP49128	zinc finger protei	758	5	3.7	11	19	AAW63537	Salvage receptor b
686	5	3.7	7	23	ABP50504	zinc finger protei	759	5	3.7	11	19	AAW59349	Mutine MIM24 antib
687	5	3.7	8	16	AAW87293	Plasmodium falci	760	5	3.7	11	19	AAW44819	Human Fab or F(ab'
688	5	3.7	8	18	AAW17459	Protein kinase C-e	761	5	3.7	11	19	AAW40573	Salvage receptor b
689	5	3.7	8	19	AAW54881	Isozyme-specific a	762	5	3.7	11	20	AAV39356	Human IgG derived
690	5	3.7	8	20	AAV53414	HIV-1 gp120 protei	763	5	3.7	11	20	AAV14576	Human erib2 recept
691	5	3.7	8	20	AAV40263	Amino acid sequenc	764	5	3.7	11	20	AAV08178	Human FabY1b salva
692	5	3.7	8	20	AAV46755	HIV-derived lipope	765	5	3.7	11	20	AAW95617	Antibody-salvage r
693	5	3.7	8	21	AAV84506	Amino acid sequenc	766	5	3.7	11	20	AAW92418	Human Fab salva
694	5	3.7	8	22	ABP11960	HIV A02 super moti	767	5	3.7	11	21	AAE26995	Salvage receptor b
695	5	3.7	8	22	ABP11961	HIV A02 super moti	768	5	3.7	11	21	AAE82352	Salvage receptor b
696	5	3.7	8	22	ABP13324	HIV A02 super moti	769	5	3.7	11	22	ABP12422	Salvage receptor b
697	5	3.7	8	22	ABP13326	HIV A02 super moti	770	5	3.7	11	22	ABP12423	HIV A02 super moti
698	5	3.7	8	22	ABP14262	HIV A03 super moti	771	5	3.7	11	22	ABP12424	HIV A02 super moti
699	5	3.7	8	22	ABP15128	HIV A24 super moti	772	5	3.7	11	22	ABP13880	HIV A03 super moti
700	5	3.7	8	22	ABP18694	HIV B62 super moti	773	5	3.7	11	22	ABP14928	HIV A03 super moti
701	5	3.7	8	22	ABP20302	HIV A03 motif env	774	5	3.7	11	22	ABP15439	HIV A24 super moti
702	5	3.7	8	22	ABP22601	HIV A11 motif env	775	5	3.7	11	22	ABP18687	HIV B62 super moti
703	5	3.7	8	22	ABP67061	Recombinant adenov	776	5	3.7	11	22	AAE12054	Salvage receptor b
704	5	3.7	8	23	AAW47749	SIV Tat(28-35). S	777	5	3.7	11	22	AAW62076	Salvage receptor b
705	5	3.7	8	23	AAW47751	CTL epitope Tat_SL	778	5	3.7	11	22	AAW66779	Salvage receptor b
706	5	3.7	8	23	AAW47758	CTL epitope 94004	779	5	3.7	11	23	ABB80871	Salvage receptor b
707	5	3.7	9	15	AAW57747	Peptide (247) inh1	780	5	3.7	11	23	ABP61196	Salvage receptor b
708	5	3.7	9	17	AAW49332	Human leucocyte an	781	5	3.7	11	23	AAV79273	IgG molecule Fe re
709	5	3.7	9	17	AAW93071	Human cartilage g1	782	5	3.7	11	23	AAE20206	Human cartilage g1
710	5	3.7	9	17	AAW95072	Human cartilage g1	783	5	3.7	11	23	AAE20212	Human cartilage g1
711	5	3.7	9	21	AAV69326	Peptide fragment o	784	5	3.7	12	12	AAE20212	Human cartilage g1
712	5	3.7	9	21	AAV69327	Human chondrocyte	785	5	3.7	12	14	AAE11562	Native HIV gp160 p
713	5	3.7	9	21	AAV77399	HIV A02 super moti	786	5	3.7	12	14	AAE83197	GP160-CR-1 amino a
714	5	3.7	9	22	ABP12125	HIV A02 super moti	787	5	3.7	12	19	AAW76962	Fusion Immunoglobu
715	5	3.7	9	22	ABP13422	HIV B58 super moti	788	5	3.7	12	21	AAW08236	Amino acid sequenc
716	5	3.7	9	22	ABP17771	Human peptide #195	789	5	3.7	12	22	AAW64141	Human Clara cell p
717	5	3.7	9	22	AAW98677	Human peptide SEQ ID	790	5	3.7	12	23	ABJ01192	Human Brn-5 transc
718	5	3.7	9	22	AAW22208	Peptide from B4-5	791	5	3.7	12	23	AAE20208	Human cartilage g1
719	5	3.7	9	22	AAW63159	Granulocyte-colony	792	5	3.7	12	23	AAU87872	p27 domain binding
720	5	3.7	9	23	AAU93454	Human cartilage g1	793	5	3.7	13	10	AAW90868	Proposed T cell ep
721	5	3.7	9	23	AAE20213	Human 34P3D7 HLA A	794	5	3.7	13	14	AAE11574	Native HIV gp160 p
722	5	3.7	9	23	ABB51446	Human 34P3D7 HLA A	795	5	3.7	13	14	AAE11574	Anti-CD18 Fab H52
723	5	3.7	9	23	ABB51455	Human 34P3D7 HLA A	796	5	3.7	13	17	AAW066901	Peptide contg. hc
724	5	3.7	9	23	ABB51673	Human 34P3D7 HLA A	797	5	3.7	13	17	AAW95075	Peptide contg. hc
725	5	3.7	9	23	ABB51756	Human 34P3D7 HLA A	798	5	3.7	13	17	AAW95076	Fusion Immunoglobu
726	5	3.7	9	23	ABB51777	Human 34P3D7 HLA A	799	5	3.7	13	19	AAW76894	Human cartilage gp
727	5	3.7	9	23	ABB51850	Human 34P3D7 HLA B	799	5	3.7	13	19	AAW61597	Human Fab variant
728	5	3.7	9	23	ABB51867	Human 34P3D7 HLA B	800	5	3.7	13	19	AAW40583	Human anti-CD18 an
729	5	3.7	9	23	ABB51971	Human 34P3D7 HLA B	801	5	3.7	13	20	AAW92429	Anti-CD18 Fab H52
730	5	3.7	9	23	ABB51976	Fusion immunoglobu	802	5	3.7	13	21	AAW27005	GP38x-derived cell
731	5	3.7	10	19	AAW76922	CTLA-4 VLD CDR3 re	803	5	3.7	13	21	AAW03439	HIV-1 envelope pep
732	5	3.7	10	20	AAV39586	HIV A02 super moti	804	5	3.7	13	21	AAW83452	HIV-1 envelope pep
733	5	3.7	10	22	ABP12280	HIV A03 super moti	805	5	3.7	13	21	AAW83453	HIV-1 envelope pep
734	5	3.7	10	22	ABP13657	HIV A03 super moti	806	5	3.7	13	21	AAW83455	HIV-1 envelope pep
735	5	3.7	10	22	ABP14843	HIV A03 motif env	807	5	3.7	13	21	AAW83455	Peptide fragment o
736	5	3.7	10	22	ABP20300	HIV A03 motif pol	808	5	3.7	13	21	AAV69230	Human chondrocyte
737	5	3.7	10	22	ABP21079	HIV A11 motif env	809	5	3.7	13	21	AAV69231	Human cartilage gp
738	5	3.7	10	22	ABP21303	HIV A11 motif pol	810	5	3.7	13	21	AAV77400	Human cartilage gp
739	5	3.7	10	22	ABP22600	HIV A11 motif pol	811	5	3.7	13	22	AAU03045	Human cartilage gp
740	5	3.7	10	22	ABP23114	HIV A11 motif pol	812	5	3.7	13	22	AAU03045	Human cartilage gp
							813	5	3.7	13	22	AAU03046	Human cartilage gp

814	5	3.7	13	22	AAU03047	Human cartilage gp	887	5	3.7	20	19	AAW23852	HIV-1 strain MN pe
815	5	3.7	13	22	AAU03048	Human cartilage gp	888	5	3.7	20	20	AAV05352	HIV-1 Cluvac pepti
816	5	3.7	13	22	AAU03049	Human cartilage gp	889	5	3.7	20	20	AAV05352	HIV-1 Cluvac pepti
817	5	3.7	13	22	AAU03050	Human cartilage gp	890	5	3.7	20	21	AAW74257	HJ loop peptide C-
818	5	3.7	13	22	AAU03051	Human cartilage gp	891	5	3.7	20	21	AAV9168	HIV class II bindi
819	5	3.7	13	22	AAU03052	Human cartilage gp	892	5	3.7	20	21	AAV78125	Treponema pallidum
820	5	3.7	13	22	AAU03053	Human cartilage gp	893	5	3.7	20	21	AAV85806	HIV-1 ORF K8.1-derl
821	5	3.7	13	22	AAU03054	Human cartilage gp	894	5	3.7	21	10	AAV90086	Peptide active aga
822	5	3.7	13	22	AAU03055	Human cartilage gp	895	5	3.7	21	14	AAV93859	E2 peptide E2-7, r
823	5	3.7	13	22	AAU03056	Human cartilage gp	896	5	3.7	21	15	AAV6479	Consensus sequence
824	5	3.7	13	22	AAU03057	Human cartilage gp	897	5	3.7	21	22	ABR31180	Peptide #3831 enco
825	5	3.7	13	22	AAU03058	Human cartilage gp	898	5	3.7	21	22	ABR36377	Peptide #3883 enco
826	5	3.7	13	22	AAU03059	Human cartilage gp	899	5	3.7	21	22	ABR21731	Protein #3730 enco
827	5	3.7	13	22	AAU03060	Human cartilage gp	900	5	3.7	21	22	AAV57142	Human brain expres
828	5	3.7	13	22	AAU03061	Human cartilage gp	901	5	3.7	21	22	AAV69535	Human bone marrow
829	5	3.7	13	22	AAU03062	Human cartilage gp	902	5	3.7	21	22	AAV17365	Peptide #3799 enco
830	5	3.7	13	22	AAU03063	Human cartilage gp	903	5	3.7	21	22	AAV29875	Peptide #3912 enco
831	5	3.7	13	22	AAU03064	Human cartilage gp	904	5	3.7	21	22	AAV05054	Peptide #3736 enco
832	5	3.7	13	22	AAU03065	Human cartilage gp	905	5	3.7	21	22	AAV88709	Human interleukin-
833	5	3.7	13	22	AAU03066	Human cartilage gp	906	5	3.7	21	23	ABG39160	Human peptide enco
834	5	3.7	13	22	AAU03067	Human cartilage gp	907	5	3.7	22	18	AAV26448	Erythropoietin rec
835	5	3.7	13	22	AAU03068	Human cartilage gp	908	5	3.7	22	22	AAV83304	Lipo peptide BP1-23
836	5	3.7	13	22	AAU03069	Human cartilage gp	909	5	3.7	23	19	AAV10886	Signal peptide for
837	5	3.7	13	22	AAU03070	Human cartilage gp	910	5	3.7	23	20	AAV653139	Amino acid sequenc
838	5	3.7	13	22	AAU03071	Human cartilage gp	911	5	3.7	23	22	AAV10886	Peptide from B4-5
839	5	3.7	13	22	AAU03072	Human cartilage gp	912	5	3.7	23	23	AAV00940	VH ligand-binding
840	5	3.7	13	22	AAU03073	Human cartilage gp	913	5	3.7	24	14	AAV26451	Human ovarian anti
841	5	3.7	13	22	AAU03074	Human cartilage gp	914	5	3.7	24	18	AAV26451	Lipo peptide TPE2-7
842	5	3.7	13	22	AAU03075	Human cartilage gp	915	5	3.7	24	15	AAV86234	Erythropoietin rec
843	5	3.7	13	22	AAU03076	Human cartilage gp	916	5	3.7	25	15	AAV86234	Oestrogen receptor
844	5	3.7	13	22	AAU03077	Human cartilage gp	917	5	3.7	25	15	AAV52775	Murine BIR-3 immu
845	5	3.7	13	22	AAU03078	Human cartilage gp	918	5	3.7	25	15	AAV52775	Murine BIR-3 immu
846	5	3.7	13	22	AAU03079	Human cartilage gp	919	5	3.7	25	22	ABR39790	Peptide #7296 enco
847	5	3.7	13	22	AAU03080	Human cartilage gp	920	5	3.7	25	22	ABR24408	Protein #407 enco
848	5	3.7	13	22	AAU03081	Human cartilage gp	921	5	3.7	25	22	AAV60511	Human brain expres
849	5	3.7	13	22	AAU03082	Human cartilage gp	922	5	3.7	25	22	AAV73164	Human bone marrow
850	5	3.7	13	22	AAU03083	Human cartilage gp	923	5	3.7	25	22	AAV33377	Peptide #7414 enco
851	5	3.7	13	22	AAU03084	Human cartilage gp	924	5	3.7	25	23	ABG43011	Human peptide enco
852	5	3.7	13	22	AAU03085	Human cartilage gp	925	5	3.7	26	15	AAV63813	HIV-1 gp120-8 epit
853	5	3.7	13	22	AAU03086	Human cartilage gp	926	5	3.7	26	22	ABR38617	Peptide #6123 enco
854	5	3.7	13	22	AAU03087	Human cartilage gp	927	5	3.7	26	22	ABR42837	Peptide #10343 enc
855	5	3.7	13	22	AAU03088	Human cartilage gp	928	5	3.7	26	22	ABR17860	Human nervous syst
856	5	3.7	13	22	AAU03089	Human cartilage gp	929	5	3.7	26	22	ABR26126	Human brain expres
857	5	3.7	13	22	AAU03090	Human cartilage gp	930	5	3.7	26	22	AAV63729	Human bone marrow
858	5	3.7	13	22	AAU03091	Human cartilage gp	931	5	3.7	26	22	AAV71786	Human bone marrow
859	5	3.7	13	22	AAU03092	Human cartilage gp	932	5	3.7	26	22	AAV6543	Human bone marrow
860	5	3.7	13	22	AAU03093	Human cartilage gp	933	5	3.7	26	22	AAV20931	Human bone marrow
861	5	3.7	13	22	AAU03094	Human cartilage gp	934	5	3.7	26	22	AAV32075	Peptide #7365 enco
862	5	3.7	13	22	AAU03095	Human cartilage gp	935	5	3.7	26	22	AAV36650	Peptide #6112 enco
863	5	3.7	13	22	AAU03096	Human cartilage gp	936	5	3.7	26	23	ABG41600	Peptide #10687 enc
864	5	3.7	13	22	AAU03097	Human cartilage gp	937	5	3.7	26	23	ABG45796	Human peptide enco
865	5	3.7	13	22	AAU03098	Human cartilage gp	938	5	3.7	27	20	AAV32945	Human peptide enco
866	5	3.7	13	22	AAU03099	Human cartilage gp	939	5	3.7	27	20	AAV05709	Mutant TD regulato
867	5	3.7	13	22	AAU03100	Human cartilage gp	940	5	3.7	27	22	ABR11917	Theonine dehydrat
868	5	3.7	13	22	AAU03101	Human cartilage gp	941	5	3.7	27	23	ABG47087	Human secreted pro
869	5	3.7	13	22	AAU03102	Human cartilage gp	942	5	3.7	28	13	AAV27308	Human peptide enco
870	5	3.7	13	22	AAU03103	Human cartilage gp	943	5	3.7	28	20	AAV41547	Peptide corresp. t
871	5	3.7	13	22	AAU03104	Human cartilage gp	944	5	3.7	28	22	ABG09553	Fragment of human
872	5	3.7	13	22	AAU03105	Human cartilage gp	945	5	3.7	28	22	ABR17359	Novel human diagno
873	5	3.7	13	22	AAU03106	Human cartilage gp	946	5	3.7	28	22	AAV86330	Human nervous syst
874	5	3.7	13	22	AAU03107	Human cartilage gp	947	5	3.7	28	22	AAV61520	Human immune/haema
875	5	3.7	13	22	AAU03108	Human cartilage gp	948	5	3.7	29	16	AAV28289	Human prostate cata
876	5	3.7	13	22	AAU03109	Human cartilage gp	949	5	3.7	29	21	AAV72943	E. coli pTase-bet
877	5	3.7	13	22	AAU03110	Human cartilage gp	950	5	3.7	29	21	AAV69225	HIV-1 non-subtype
878	5	3.7	13	22	AAU03111	Human cartilage gp	951	5	3.7	29	12	ABR79266	Human prostate spe
879	5	3.7	13	22	AAU03112	Human cartilage gp	952	5	3.7	30	13	AAV41244	HIV gp120 epitope
880	5	3.7	13	22	AAU03113	Human cartilage gp	953	5	3.7	30	18	AAV25746	Partial human SIM
881	5	3.7	13	22	AAU03114	Human cartilage gp	954	5	3.7	30	13	AAV30759	Novel human diagno
882	5	3.7	13	22	AAU03115	Human cartilage gp	955	5	3.7	30	22	ABG21728	Staphylococcus hyl
883	5	3.7	13	22	AAU03116	Human cartilage gp	956	5	3.7	30	22	AAV51585	Peptide #10509 enc
884	5	3.7	13	22	AAU03117	Human cartilage gp	957	5	3.7	30	22	ABR43003	Human brain expres
885	5	3.7	13	22	AAU03118	Human cartilage gp	958	5	3.7	30	22	AAV63911	Human bone marrow
886	5	3.7	13	22	AAU03119	Human cartilage gp	959	5	3.7	30	22	AAV76727	Peptide #10871 enc
887	5	3.7	13	22	AAU03120	Human cartilage gp	960	5	3.7	30	22	AAV36834	

```

960 5 3.7 30 23 ABG45918 Human peptide enco
961 5 3.7 30 23 AA084541 HIV ENV segment 7.
962 5 3.7 30 23 AA084542 HIV ENV segment 8.
963 5 3.7 31 10 AAP93048 HIV env protein an
964 5 3.7 31 17 AAM16945 Random recombinant
965 5 3.7 31 18 AAM25508 Random peptide rec
966 5 3.7 31 22 ABB32432 Peptide #5083 enco
967 5 3.7 32 10 AAP93050 HIV env protein an
968 5 3.7 32 10 AAB61562 HIV env protein an
969 5 3.7 33 10 AAP93047 HIV env protein an
970 5 3.7 33 16 AAR87198 Heat shock protein
971 5 3.7 33 22 ABR42193 Peptide #9699 enco
972 5 3.7 33 22 ABB16041 Human nervous syst
973 5 3.7 33 22 ABB25743 Protein #7742 enco
974 5 3.7 33 22 AAM63078 Human brain expres
975 5 3.7 33 22 AAM75889 Human bone marrow
976 5 3.7 33 22 AAM35647 Peptide #9684 enco
977 5 3.7 33 22 AAM36000 Peptide #10037 enc
978 5 3.7 33 23 ABG45332 Human peptide enco
979 5 3.7 34 20 AAY39495 HCV E1 protein res
980 5 3.7 34 20 AAY14167 HCV envelope regio
981 5 3.7 34 20 AAY12918 Amino acid sequenc
982 5 3.7 34 21 AAG55417 Arabidopsis thalia
983 5 3.7 34 21 AAY69657 HCV E1 peptide. v3
984 5 3.7 34 22 AAB60939 Horseshoe crab rec
985 5 3.7 35 9 AAP82462 Peptide component
986 5 3.7 35 16 AAR66434 PCLUS 1-18 (102-12
987 5 3.7 35 16 AAR66408 PCLUS 1-18 (102-12
988 5 3.7 35 20 AAY25469 Insulin-like growt
989 5 3.7 35 20 AAY05345 HIV-1 CLUVAAC pepit
990 5 3.7 35 20 AAY05338 HIV-1 CLUVAAC pepit
991 5 3.7 35 20 AAY13047 Human secreted pro
992 5 3.7 35 22 AAB71912 C. glutamicum meta
993 5 3.7 35 22 AABG09255 Novel human diagno
994 5 3.7 35 22 AAE10943 Mouse masel protei
995 5 3.7 35 22 AAM61166 Human immune/haema
996 5 3.7 35 22 AAB79744 Corynebacterium g1
997 5 3.7 35 23 ABB68640 HIV-1 gp160 V3 100
998 5 3.7 35 23 ABB68647 HIV-1 gp160 V3 100
999 5 3.7 36 20 AAY41304 HIV-1 gp120 region
1000 5 3.7 36 20 AAY14168 HCV envelope regio

```

ALIGNMENTS

```

RESULT 1
AAB60646 standard; Protein: 134 AA.
AC
XX AAB60646;
XX
XX 04-MAY-2001 (first entry)
XX
DE Moraxella catarrhalis strain ATCC43617 BASB125 protein.
XX
XX BASB125 protein; strain ATCC43617; antigen: antibody; vaccine;
XX genetic immunisation; infection; upper respiratory tract; otitis media;
XX hearing loss; deafness; pneumonia; sinusitis; nosocomial infection;
XX invasive disease; antibacterial; auditory.
XX
XX Moraxella catarrhalis.
XX
XX WO200109331-A2.
XX
XX 08-FEB-2001.
XX
XX 27-JUL-2000; 2000WO-EP07291.
XX
XX 30-JUL-1999; 99GB-0018041.
XX
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX

```

```

PI Thonnard J;
XX WPI; 2001-168707/17.
DR N-PSDB; AAF59800.
XX
PT New BASB125 polypeptide isolated from Moraxella catarrhalis for
PT treating, preventing and diagnosing diseases associated with M.
PT catarrhalis infection in mammals, e.g. otitis media in humans
XX
XX Claim 4; Page 64; 73pp: English.
XX
CC The invention relates to the Moraxella catarrhalis strain ATCC43617
CC BASB125 protein (AAB60646) and to DNA encoding it (AAF59800). The
CC invention also relates to immunogenic fragments of the BASB125 protein,
CC expression vectors and host cells comprising BASB125 nucleic acids, the
CC recombinant production of BASB125, vaccine compositions comprising the
CC BASB125 protein or nucleic acid, an antibody against BASB125, therapeutic
CC compositions comprising the anti-BASB125 antibody, and a method of
CC identifying a Moraxella catarrhalis infection via the detection of
CC BASB125 proteins or antibodies. The vaccine compositions of the invention
CC are useful as prophylactic or therapeutic agents against Moraxella
CC catarrhalis infections in mammals, particularly humans. Moraxella
CC catarrhalis is a Gram negative bacterium frequently isolated from the
CC human upper respiratory tract, which is responsible for several
CC pathological conditions. It is responsible for about 13% of otitis media
CC cases in children (which can lead to temporary or permanent hearing
CC loss). It also causes pneumonia in elderly people, and sinusitis,
CC nosocomial infections and, less frequently, invasive diseases. BASB125
CC proteins or nucleotides may additionally be used in screening for novel
CC antibacterial compounds, and in the diagnosis and staging of infections.
CC ATCC43617 BASB125 protein.
XX
XX Sequence 134 AA:
XX
XX Query Match 100.0%; Score 134; DB 22; Length 134;
XX Best Local Similarity 100.0%; Pred. No. 3.9e-133;
XX Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 MKLLTYATATMTAFTLASCASPESNPKNSSANLTLTKHAKONCOROLGHOYWKI 60
DB 1 MKLLTYATATMTAFTLASCASPESNPKNSSANLTLTKHAKONCOROLGHOYWKI 60
OY 61 AAMKLSSESRAKISSETACGCAADKAPAVSLTETTTAIPNATEVAOKIVRHSKPCM 120
DB 61 AAMKLSSESRAKISSETACGCAADKAPAVSLTETTTAIPNATEVAOKIVRHSKPCM 120
OY 121 LETYNATIVPTTTR 134
DB 121 LETYNATIVPTTTR 134
XX
XX RESULT 2
XX AAY89752 standard; peptide: 32 AA.
XX
XX AAY89752;
XX
XX 23-MAY-2000 (first entry)
XX
XX Core polypeptide fragment T No. 1317.
XX
XX Retrovirus; hybrid polypeptide; enhancer; gp41; envelope protein; HIV-1;
XX HIV-2; SIV; pharmacokinetic; half-life; growth factor; cytokine; viral;
XX anti-fusogenic; differentiation factor; interleukin; interferon;
XX colony stimulating factor; hormone; angiogenic factor.
XX
XX unidentified.
XX
XX WO9959615-A1.
XX
XX 25-NOV-1999.
XX
XX
XX
XX

```

PF 20-MAY-1999; 99WO-US11219.
XX
PR 20-MAY-1998; 98US-0082279.
XX
PA (TRIM-) TRIMERIS INC.
XX
PI Barney S, Guthrie KI, Merutka G, Anwer MK, Lambert DM;
XX
DR WPI; 2000-136792/12.
XX
PT A new hybrid polypeptide with enhanced pharmacokinetic properties
XX comprises enhancer sequence -
XX
PS Disclosure; Page 43; 124pp; English.
XX
CC The invention relates to hybrid polypeptides comprising enhancer peptide
CC sequence linked to core polypeptides. The enhancer polypeptides are
CC derived from various retroviral envelope (gp41) protein sequences,
CC especially from HIV-1, HIV-2 and SIV. The enhancer peptides enhance the
CC pharmacokinetic properties such as increasing the half-life of any core
CC polypeptide that they are linked to. The core polypeptides are any
CC polypeptide that may be introduced into a living system and that can
CC function as a pharmacologically useful peptide for the treatment or
CC prevention of a disease. The core polypeptides are bioactive peptides
CC selected from a growth factor, cytokine, differentiation factor,
CC interleukin, interferon, colony stimulating factor, hormone or
CC angiogenic factor. The peptides of the invention can be used for
CC inhibiting viral infection and can be used in anti-viral and
CC anti-fusogenic treatments. Sequences AY88651-Y90055 represent core
CC polypeptide fragments that can be used in the invention. Some sequences
CC among those indicated also comprise enhancer fragments at terminal ends
XX and form hybrid polypeptides.
XX
SQ Sequence 32 AA:

Query Match 5.2%; Score 7; DB 21; Length 32;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 33 ANLTSL 39
DB 6 ANLTSL 12

RESULT 3
ID ABB01160 standard; Peptide: 32 AA.
XX
AC ABB01160;
XX
DT 03-JAN-2002 (first entry)
XX
DE Viral DP178/107-like region peptide T1317.
XX
XX Human immunodeficiency virus; HIV; respiratory syncytial virus; RSV;
XX virucide; heptad repeat region; transmembrane protein; gp41; HRL; HR2;
XX infection.
XX
OS Viridae.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal is substituted by Ac"
FT Modified-site 32 /note= "C-terminal amide"
FT
XX
XX WO200164013-A2.
XX
XX 07-SEP-2001.
XX
XX 07-FEB-2001; 2001WO-US03988.
XX
XX 29-FEB-2000; 2000US-0515965.
XX
PR

XX
XX (TRIM-) TRIMERIS INC.
XX
PI Antczak JB, Delmedico MK, Erickson JB, Lambert DM, Sista P;
XX
DR WPI; 2001-514829/56.
XX
XX Heptad repeat region peptide analogs useful for inhibiting virus/cells
XX fusion, useful for treating HIV and Respiratory Syncytial Virus
XX infection -
XX
PS Disclosure; Page 56; 587pp; English.
XX

CC The invention relates to isolated analogues of the heptad repeat region
CC peptides DP178 and DP107. DP178 and DP107 correspond to amino acids
CC 638-673 (heptad repeat region HR2) and 558-595 (heptad repeat region
CC HR1) respectively, of HIV-1LAI transmembrane protein gp41. The HRL
CC and HR2 regions of proteins interact non-covalently with each other
CC and/or with peptides derived from them. This interaction is required for
CC normal infectivity of viruses such as RSV and HIV. The heptad
CC repeat region peptide analogues may be used to inhibit respiratory
CC syncytial virus (RSV) infection in a cell. They may also be used to
CC inhibit HIV infection. The present sequence is a peptide provided in
XX the specification.
XX
SQ Sequence 32 AA:

Query Match 5.2%; Score 7; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 33 ANLTSL 39
DB 6 ANLTSL 12

RESULT 4
ID AAU13706 standard; Peptide: 32 AA.
XX
AC AAU13706;
XX
DT 21-NOV-2001 (first entry)
XX
DE DP178-like/DP107-like peptide T-1317.
XX
XX Anti-retroviral; DP178-like; DP107-like; transmembrane protein gp41;
XX antifusogenic; antiviral; HIV transmission; mutant; mutein.
XX
OS Human immunodeficiency virus 1 isolate LAI.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal is substituted by Ac"
FT Modified-site 32 /note= "C-terminal amide"
FT
XX
XX WO200151673-A2.
XX
XX 19-JUL-2001.
XX
XX 05-JUL-2000; 2000WO-US35727.
XX
XX 09-JUL-1999; 99US-0350841.
XX
XX (TRIM-) TRIMERIS INC.
XX
XX Jeffs P, Lackey JM, Erickson JB, Lawless MK, Merutka G;
XX
XX WPI; 2001-442157/47.
XX
XX Identifying a compound that inhibits the formation of or disrupts a
XX
XX

PT DP107/DP178 complex, especially compounds with antifeosgenic, antiviral
PT or intracellular modulatory activity, by detecting the formation of a
PS DP107/DP178 complex -
XX Disclosure: Page 76; 259pp; English.

CC The present invention relates to peptides which exhibit anti-retroviral
CC activity. The peptides of the invention (AAU12559-AAU14009) comprise
CC DP178-like and DP107-like peptides. The DP178 peptide corresponds
CC to amino acids 639-673 of the transmembrane protein gp41 from human
CC immunodeficiency virus 1 (HIV-1) isolate LAI. The DP107 peptide
CC corresponds to amino acids 558-595 of gp41 from HIV-1LAI. The invention
CC also relates to a method of identifying compounds that inhibit the
CC formation of or disrupts a DP107/DP178 complex. The method comprises
CC detecting the formation of a DP107/DP178 complex, both in the presence
CC or absence of a test compound, in a reaction mixture containing DP107
CC and DP178 peptides. The method is useful for identifying compounds,
CC including small molecule compounds, which may themselves exhibit
CC antifeosgenic, antiviral or intracellular modulatory activity. The
CC DP178-like/DP107-like peptides are useful to inhibit human and non-human
CC retroviral, particularly HIV, transmission to uninfected cells. The
CC present sequence represents one of the DP178-like/DP107-like peptides
CC of the invention.

SO Sequence 32 AA;

Query Match 5.2%; Score 7; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39
Db 6 ANLTSL 12

RESULT 5
AAB78153
ID AAB78153 standard; Peptide: 32 AA.
XX
AC AAB78153;
XX
DT 19-APR-2001 (first entry)
XX
DE Core polypeptide T1317.
XX
XX Core polypeptide; enhancer; antiviral; anti-HIV;
XX virucide; hepatotropic; antiinflammatory; hybrid polypeptide;
XX coiled-coil peptide interaction; fusion-related disorder;
XX bacterial infection; viral infection.
XX
OS Unidentified.
XX
XX WO200103723-A1.
XX
XX 18-JAN-2001.
XX
XX 10-JUL-2000; 2000WO-US18772.
XX
XX 09-JUL-1999; 99US-0350641.
XX
XX (TRIM-) TRIMERIS INC.
XX
XX Barney S, Guthrie KI, Merutka G, Anwer MK, Lambert DM;
XX WPI; 2001-147136/15.
XX
XX New hybrid polypeptide, useful for preventing, treating and diagnosing
XX e.g. viral infections, comprises an enhancer peptide linked to a core
XX polypeptide -
XX
XX Disclosure: Page 56; 151pp; English.
XX
XX The present sequence is a core polypeptide which may be linked to

CC an enhancer peptide to form a novel hybrid polypeptide. The hybrid
CC polypeptide exhibits enhanced pharmacokinetic properties relative to
CC those exhibited by the core polypeptide when introduced into a living
CC system. It is used to increase the in vitro or ex vivo half-life of
CC the core polypeptide. The hybrid and core polypeptides can be used for
CC modulating fusogenic events and intracellular processes involving
CC coiled-coil peptide interactions. Other uses include preventing,
CC treating and/or diagnosing disorders involving fusion events (e.g.
CC modulation of neurotransmitter exchange and sperm-egg fusion),
CC intracellular processes involving coiled-coil peptides (e.g. bacterial
CC infections) and viral infections that involve cell-cell and/or
CC virus-cell fusion (e.g. viral infections caused by human
CC immunodeficiency virus, respiratory syncytial virus, Epstein-Barr
CC virus, hepatitis B virus, Mason-Pfizer virus and polio virus).
CC The enhancer peptide sequence increases the half-life and reduces the
CC clearance rate of therapeutic peptides, which increases their efficacy
CC and minimises the incidence and severity of adverse side effects.
CC In addition, this increases the sensitivity of the diagnostic procedure
CC in which they are used.

SO Sequence 32 AA;

Query Match 5.2%; Score 7; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39
Db 6 ANLTSL 12

RESULT 6
ABB02611
ID ABB02611 standard; Peptide: 40 AA.
XX
XX ABB02611;
XX
DT 03-JAN-2002 (first entry)
XX
XX Viral core polypeptide, SEQ ID NO: 1138.
XX
XX Human immunodeficiency virus; HIV; respiratory syncytial virus; RSV;
XX virucide; heptad repeat region; transmembrane protein; gp41; HR1; HR2;
XX infection.
XX
XX Virididae.
XX
XX WO200164013-A2.
XX
XX 07-SEP-2001.
XX
XX 07-FEB-2001; 2001WO-US03988.
XX
XX 29-FEB-2000; 2000US-0515965.
XX
XX (TRIM-) TRIMERIS INC.
XX
XX Antczak JB, Delmedico MK, Erickson JB, Lambert DM, Sista P;
XX WPI; 2001-514829/56.
XX
XX Heptad repeat region peptide analogs useful for inhibiting virus/cells
XX fusion, useful for treating HIV and respiratory syncytial virus
XX infection -
XX
XX Disclosure: Page 468; 587pp; English.
XX
XX The invention relates to isolated analogues of the heptad repeat region
XX peptides DP178 and DP107. DP178 and DP107 correspond to amino acids
XX 638-673 (heptad repeat region HR2) and 558-595 (heptad repeat region
XX HR1) respectively, of HIV-1LAI transmembrane protein gp41. The HR1
XX and HR2 regions of proteins interact non-covalently with each other
XX and/or with peptides derived from them. This interaction is required for

CC normal infectivity of viruses such as RSV and HIV. The heptad
CC repeat region peptide analogues may be used to inhibit respiratory
CC syncytial virus (RSV) infection in a cell. They may also be used to
CC inhibit HIV infection. The present sequence is a peptide provided in
CC the specification.

SQ Sequence 40 AA;

Query Match 5.2%; Score 7; DB 22; Length 40;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39
|||||
Db 6 ANLTSL 12

RESULT 7
AAM82470
ID AAM82470 standard; Protein: 62 AA.

AC AAM82470;

DT 07-NOV-2001 (first entry)

XX Human Immune/haematopoietic antigen SEQ ID NO:10063.

XX Human; Immune: haematopoietic; immune/haematopoietic antigen; cancer;
KW Cytostatic; gene therapy; vaccine; metastasis.

OS Homo sapiens.

XX WO200157182-A2.

XX PD 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-019874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-020515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 11-JUL-2000; 2000US-0216880.

XX 14-JUL-2000; 2000US-0217496.

XX 26-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 14-AUG-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

XX 14-AUG-2000; 2000US-0225758.

XX 18-AUG-2000; 2000US-0226279.

XX 22-AUG-2000; 2000US-0226681.

XX 22-AUG-2000; 2000US-0226868.

XX 23-AUG-2000; 2000US-0227182.

XX 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 12-SEP-2000; 2000US-0232081.

PR 14-SEP-2000; 2000US-0232397.

PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.

PR 14-SEP-2000; 2000US-0232400.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.

PR 25-SEP-2000; 2000US-0234998.

PR 26-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.

PR 02-OCT-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237039.

PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.

PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.

PR 08-NOV-2000; 2000US-0246527.

PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.

PR 08-NOV-2000; 2000US-0246539.

PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.

PR 08-NOV-2000; 2000US-0246613.

PR 17-NOV-2000; 2000US-0249207.

PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.

PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XI
 DR WPI: 2001-483426/52.
 DR N-PSDB; AAK55251.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 XX
 Claim 11: SEQ ID NO 10063; 3071bp + Sequence Listing: English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK67694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SO Sequence 62 AA:

Query Match 5.2%; Score 7; DB 22; Length 62;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 85 APEAVSL 91
 Db 55 APEAVSL 61

RESULT 8
 AAG01641

ID AAG01641 standard; Protein; 77 AA.
 XX
 AC AAG01641;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 5722.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EPI033401-A2.
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GEST) GENSET.
 XX
 PI Dunas Mline Edwards J, Duclert A, Giordano J;
 XX
 DR WPI: 2000-500381/45.
 DR N-PSDB; AAC01647.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 13: SEQ ID 5722; 71bp + CD-ROM; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX
 SO Sequence 77 AA:

Query Match 5.2%; Score 7; DB 21; Length 77;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 90 SLTELTFT 96
 Db 69 SLTELTFT 75

RESULT 9
 ABP33369
 ID ABP33369 standard; Protein; 87 AA.
 XX
 AC ABP33369;
 XX
 DT 08-JUL-2002 (first entry)

Human synthase-like ORF2342 protein, SEQ ID NO:4684.
 DE
 XX
 KW Human; ORF; open reading frame; ORF; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; hematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;

KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KW vasotropic; antiporiatic; antidiabetic; cyostatic; noctropic;
 KW neuroprotective; antithrombotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antihypertoid; antinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.
 XX
 OS Homo sapiens.
 PN WO200190366-A2.
 PD 29-NOV-2001.
 XX
 XX 24-MAY-2001: 2001WO-US17076.
 PF
 XX 24-MAY-2000: 2000US-206690P.
 PR
 XX (CURA-) CURAGEN CORP.
 PA
 XX Leach MD, Shinkets RA;
 PI
 DR WPI: 2002-106200/14.
 DR N-PSDB; ABN77395.
 XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation -
 PS
 PS Claim 10: Page 1424; 2508pp; English.
 XX
 CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antinflammatory activity, tumour inhibition activity,
 CC and antinefective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.
 XX
 SO Sequence 87 AA:

Query Match 5.2%; Score 7; DB 23; Length 87;
 Best Local Similarity 100.0%; Pred. No. 20;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 30 NSSANLT 36
 |||||
 Db 28 NSSANLT 34
 RESULT 10
 AAG01640
 ID AAG01640 standard; Protein; 131 AA.
 XX
 AC AAG01640:
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 5721.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 PD 06-SEP-2000.
 PD 21-FEB-2000: 2000EP-0200610.
 PF
 XX 26-FEB-1999: 99US-0122487.
 PR
 XX (GEST) GENSET.
 PA
 XX Dumas Mline Edwards J, Duclert A, Giordano J;
 PI
 PI WPI: 2000-500381/45.
 DR N-PSDB; AAC01646.
 DR
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 13: SEQ ID 5721: 71pp + CD-ROM; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX
 SO Sequence 131 AA:

Query Match 5.2%; Score 7; DB 21; Length 131;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 STLELT 96
 |||||
 Db 81 STLELT 87

RESULT 11
 AAG55444
 ID AAG55444 standard; Protein; 147 AA.
 AC AAG55444:

```
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 71093.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EPI03405-A2.
XX
XX 06-SEP-2000.
XX
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127452.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 18-MAY-1999; 99US-0134221.
XX 18-MAY-1999; 99US-0134370.
XX 19-MAY-1999; 99US-0134768.
XX 20-MAY-1999; 99US-0134941.
XX 21-MAY-1999; 99US-0135124.
XX 24-MAY-1999; 99US-0135353.
XX 25-MAY-1999; 99US-0135629.
XX 27-MAY-1999; 99US-0136021.
XX 28-MAY-1999; 99US-0136392.
XX 01-JUN-1999; 99US-0136782.
XX 03-JUN-1999; 99US-0137222.
XX 04-JUN-1999; 99US-0137528.
XX 07-JUN-1999; 99US-0137502.
XX 08-JUN-1999; 99US-0137724.
XX 10-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 14-JUN-1999; 99US-0138847.
XX 16-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 17-JUN-1999; 99US-0139453.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 24-JUN-1999; 99US-0140354.
XX 28-JUN-1999; 99US-0140695.
XX 29-JUN-1999; 99US-0140823.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
XX 02-JUL-1999; 99US-0142154.
XX 06-JUL-1999; 99US-0142055.
XX 08-JUL-1999; 99US-0142803.
XX 09-JUL-1999; 99US-0142920.
XX 12-JUL-1999; 99US-0142977.
XX 13-JUL-1999; 99US-0143542.
XX 14-JUL-1999; 99US-0143624.
XX 15-JUL-1999; 99US-0144005.
XX 16-JUL-1999; 99US-0144085.
XX 16-JUL-1999; 99US-0144086.
XX 19-JUL-1999; 99US-0144325.
XX 19-JUL-1999; 99US-0144331.
XX 19-JUL-1999; 99US-0144332.
XX 19-JUL-1999; 99US-0144333.
XX 19-JUL-1999; 99US-0144334.
XX 20-JUL-1999; 99US-0144335.
XX 20-JUL-1999; 99US-0144352.
XX 20-JUL-1999; 99US-0144632.
XX 21-JUL-1999; 99US-0144884.
XX 21-JUL-1999; 99US-0144814.
XX 21-JUL-1999; 99US-0145086.
XX 22-JUL-1999; 99US-0145088.
XX 22-JUL-1999; 99US-0145085.
XX 22-JUL-1999; 99US-0145087.
XX 22-JUL-1999; 99US-0145089.
XX 23-JUL-1999; 99US-0145192.
XX 23-JUL-1999; 99US-0145145.
XX 23-JUL-1999; 99US-0145218.
XX 26-JUL-1999; 99US-0145224.
XX 27-JUL-1999; 99US-0145276.
XX 27-JUL-1999; 99US-0145913.
XX 27-JUL-1999; 99US-0145918.
XX 28-JUL-1999; 99US-0145919.
XX 02-AUG-1999; 99US-0145951.
XX 02-AUG-1999; 99US-0146386.
XX 02-AUG-1999; 99US-0146388.
XX 03-AUG-1999; 99US-0146389.
XX 04-AUG-1999; 99US-0147038.
XX 04-AUG-1999; 99US-0147204.
XX 05-AUG-1999; 99US-0147302.
XX 06-AUG-1999; 99US-0147192.
XX 06-AUG-1999; 99US-0147260.
XX 06-AUG-1999; 99US-0147303.
XX 09-AUG-1999; 99US-0147416.
XX 09-AUG-1999; 99US-0147493.
XX 10-AUG-1999; 99US-0147935.
XX 11-AUG-1999; 99US-0148177.
XX 12-AUG-1999; 99US-0148319.
XX 13-AUG-1999; 99US-0148341.
XX 13-AUG-1999; 99US-0148565.
XX 16-AUG-1999; 99US-0148684.
XX 17-AUG-1999; 99US-0149368.
XX 18-AUG-1999; 99US-0149175.
XX 20-AUG-1999; 99US-0149426.
XX 20-AUG-1999; 99US-0149722.
XX 20-AUG-1999; 99US-0149723.
XX 23-AUG-1999; 99US-0149929.
XX 23-AUG-1999; 99US-0149902.
XX 25-AUG-1999; 99US-0149930.
XX 26-AUG-1999; 99US-0150566.
XX 26-AUG-1999; 99US-0150884.
```

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 5.28; Score 7; DB 21; Length 147;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 83 DKAPEAV 89
| | | | | | |
Db 34 DKAPEAV 40

RESULT 12
ABG23244
ID ABG23244 standard; Protein: 154 AA.
XX

AC ABG23244;

XX 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #23235.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
KW

XX OS Homo sapiens.
XX WO200175067-A2.
XX PN 11-OCT-2001.
XX PD 30-MAR-2001; 2001WO-US08631.
XX PF 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Dmanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
XX N-PSDB; AAS87431.
XX DR
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensic, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

Claim 20; SEQ ID No 53603; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

Sequence 154 AA;

Query Match 5.28; Score 7; DB 22; Length 154;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 87 EAVSLTE 93
| | | | | | |
Db 50 EAVSLTE 56

RESULT 13
AAG55443
ID AAG55443 standard; Protein: 157 AA.
XX

AC AAG55443;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 71092.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX

```
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126264.
PR 01-APR-1999; 99US-0126785.
PR 06-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130444.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137328.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145221.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
```

```

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 5.2%; Score 7; DB 21; Length 157;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 83 DKAPAV 89
Db 44 DKAPAV 50

```

RESULT 14

```
AAU12129
ID AAU12129 standard; Protein; 170 AA.

```

```
AAU12129;

```

```
26-MAR-2002 (first entry)

```

```
Arabidopsis RPW8.1 from strain Can.

```

```
RPW8.1; resistance to powdery mildew; Erysiphe cichoracearum;

```

```
E. cruciferarum; E. orontii; Oidium lycopersici; transgenic plant.

```

```
Arabidopsis thaliana accession Can.

```

```
WO200198479-A2.

```

```
27-DEC-2001.

```

```
19-JUN-2001; 2001WO-GB02693.

```

```
XX

```

```

PR 20-JUN-2000; 2000GB-0015122.
XX
XX (PLAN-) PLANT BIOSCIENCE LTD.
PA
XX xiao SY, Turner JG, Coleman M, Ellwood S;
PI
XX WPI: 2002-114581/15.
DR N-PSDB; AAS19498.
XX
XX Plant resistance genes from Arabidopsis thaliana and related homologues
PT useful for manipulating the resistance of plants to powdery mildew -
PS
XX
XX Claim 28; Page -: 104pp; English.

```

The invention relates to a nucleic acid molecule comprising an RPW (resistance to powdery mildew, a disease caused by Erysiphe cichoracearum, E. cruciferarum, E. orontii and Oidium lycopersici) nucleotide sequence encoding an RPW resistance polypeptide having an N-terminal transmembrane domain and a coiled coil domain and which is capable of recognising and activating in a plant into which the nucleic acid is introduced a specific defence response to challenge with a powdery mildew pathogen. Also included are primers and probes designed from conserved sequences from the cDNAs for RPW8.1 or RPW8.2, which may be used to amplify RPW8.1/8.2 sequences, a recombinant vector comprising the RPW polynucleotide, a host cell comprising the vector and a transgenic plant comprising the polynucleotide. The RPW polynucleotide is useful for influencing or affecting, in particular increasing the degree of resistance of a plant to a powdery mildew pathogens such as Erysiphe cichoracearum, E. cruciferarum, E. orontii and Oidium lycopersici. The polynucleotide, primers or complement of the polynucleotide are useful for identifying, cloning or determining the presence of RPW nucleotide sequence within a plant. The vector is useful for transforming a host cell for producing a transgenic plant such as wheat, barley, tomato or Nicotiana sp., by introducing the vector into a host cell and optionally causing or allowing recombination between the vector and the host cell genome so as to transform the host cell. The present sequence is the RPW8.1 protein from A. thaliana accession Can.

CC Note: The present sequence does not appear in the specification but CC was created by the indexer using the information on pages 85-86 and the CC RPW8.1 sequence appearing as AAU12124.

```
Sequence 170 AA;

```

```

Query Match 5.2%; Score 7; DB 23; Length 170;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 70 KAKISFT 76
Db 126 KAKISFT 132

```

RESULT 15

```
AAU12130
ID AAU12130 standard; Protein; 170 AA.

```

```
AAU12130;

```

```
26-MAR-2002 (first entry)

```

```
Arabidopsis RPW8.1 from strain Nd.

```

```
RPW8.1; resistance to powdery mildew; Erysiphe cichoracearum;

```

```
E. cruciferarum; E. orontii; Oidium lycopersici; transgenic plant.

```

```
Arabidopsis thaliana accession Nd.

```

```
WO200198479-A2.

```

```
27-DEC-2001.

```

```
19-JUN-2001; 2001WO-GB02693.

```

```
PF

```

```

XX 20-JUN-2000; 2000GB-0015122.
PR XX (PLAN-) PLANT BIOSCIENCE LTD.
PA XX
XX Xiao SY, Turner JG, Coleman M, Ellwood S;
XX WPI: 2002-114581/15.
DR N-PSDB; AAS19499.
XX
PT Plant resistance genes from Arabidopsis thaliana and related homologues
PT useful for manipulating the resistance of plants to powdery mildew -
XX
PS Claim 28; Page -: 104pp; English.
XX
CC The invention relates to a nucleic acid molecule comprising an RPW
CC (resistance to powdery mildew, a disease caused by Erysiphe
CC clhoracearum, E. cruciferarum, E. orontii and Oldium lycopersici)
CC nucleotide sequence encoding an RPW resistance polypeptide having an
CC N-terminal transmembrane domain and a coiled coil domain and which is
CC capable of recognising and activating in a plant into which the nucleic
CC acid is introduced a specific defence response to challenge with a
CC powdery mildew pathogen. Also included are primers and probes
CC designed from conserved sequences from the cDNAs for RPW8.1 or
CC RPW8.2, which may be used to amplify RPW8.1/8.2 sequences, a
CC recombinant vector comprising the RPW polynucleotide, a host cell
CC comprising the vector and a transgenic plant comprising the
CC polynucleotide. The RPW polynucleotide is useful for influencing or
CC affecting, in particular increasing the degree of resistance of a plant
CC to a powdery mildew pathogens such as Erysiphe clhoracearum, E.
CC cruciferarum, E. orontii and Oldium lycopersici. The polynucleotide,
CC primers or complement of the polynucleotide are useful for identifying,
CC cloning or determining the presence of RPW nucleotide sequence within a
CC plant. The vector is useful for transforming a host cell for producing a
CC transgenic plant such as wheat, barley, tomato or Nicotiana sp., by
CC introducing the vector into a host cell and optionally causing or
CC allowing recombination between the vector and the host cell genome so as
CC from A. thaliana accession Nd.
CC Note: The present sequence does not appear in the specification but
CC was created by the indexer using the information on pages 85-86 and the
CC RPW8.1 sequence appearing as AMU12124.
XX
SQ Sequence 170 AA:
Query Match 5.2%; Score 7; DB 23; Length 170;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 70 KAKIST 76
DB 126 KAKIST 132

```

```

PD 10-JUN-1999.
XX
XX 27-NOV-1998; 98WO-IB01939.
PF
XX 04-NOV-1998; 98US-0107077.
PR 28-NOV-1997; 97ER-0015041.
PR 17-DEC-1997; 97FR-0016034.
XX
PA (GEST ) GENSET.
XX
XX Griffiths R;
XX
DR WPI: 1999-371125/31.
XX
PT Genome sequence of Chlamydia trachomatis
PT
XX Disclosure; Page 1023-1024; 1755pp; English.
PS
XX
CC AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
CC of Chlamydia trachomatis (see AA201425). The polypeptides can be used as
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC can also be used to control growth of the microorganism. Chlamydia
CC trachomatis is responsible for a large number of diseases, e.g. eye
CC diseases such as conventional trachoma, nonendemic trachoma,
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
CC perihapatitis, Bartholinitis; pneumonia; lymphogranulomatosis; The polypeptides of the invention
CC may be of use in treating these diseases.
XX
SQ Sequence 205 AA:
Query Match 5.2%; Score 7; DB 20; Length 205;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 ANLTSL 39
DB 168 ANLTSL 174

```

```

RESULT 17
ABB60422
ID ABB60422 standard; Protein; 215 AA.
XX
XX ABB60422;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 8058.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
OS Drosophila melanogaster.
XX
XX WO200171042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
PR
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI: 2001-656860/75.
DR
DR N-PSDB; ABL04525.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT

```

PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure: SEQ ID NO 8058; 21np + Sequence Listing: English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB57737-AB72012).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 215 AA:

Query Match 5.2%; Score 7; DB 22; Length 215;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 NPMARTE 106
|111111|
DB 107 NPMARTE 113

RESULT 18
AAU35475
ID AAU35475 standard; Protein: 345 AA.
XX
XX AAU35475;
AC
XX
DT 14-FEB-2002 (first entry)
XX
DE Haemophilus influenzae cellular proliferation protein #116.
XX
KM Antisense: prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Haemophilus influenzae.
XX
XX WO200170955-A2.
PN
XX
PD 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09180.
PF
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI: 2001-611495/70.
DR N-PSDB: AMS53334.
XX
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 11068; 511np; English.
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential

CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 345 AA:

Query Match 5.2%; Score 7; DB 22; Length 345;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 SESKAKI 73
|111111|
DB 111 SESKAKI 117

RESULT 19
AAG24041
ID AAG24041 standard; Protein: 435 AA.
XX
XX AAG24041;
AC
XX
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 27566.
DE
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142054.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149729.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 5.28; Score 7; DB 21; Length 435;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 TAFTLAS 19
DB 13 TAFTLAS 19

RESULT 20
AAG24040
ID AAG24040 standard; Protein; 440 AA.
XX
AC AAG24040;
XX
XX 17-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana protein fragment SEQ ID NO: 27565.
DE
XX Protein Identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
PN
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142970.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match
Best Local Similarity 5.2%; Score 7; DB 21; Length 440;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TARTLAS 19
DB 18 TARTLAS 24

RESULT 21
AAG36664
ID AAG36664 standard; Protein; 466 AA.
XX
AC AAG36664;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 44964.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hydridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0126785.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128274.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150356.
PR 25-AUG-1999; 99US-0150356.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156556.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 5.2%; Score 7; DB 21; Length 466;
Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 64 KLSSESK 70
|||||
Db 379 KLSSESK 385

RESULT 22
AAG39612
ID AAG39612 standard; Protein; 466 AA.
XX
AC AAG39612;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49038.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127452.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132485.
PR 05-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144652.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145513.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

Query Match 5.2%; Score 7; DB 21; Length 466;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 64 KLSSESK 70
DB 379 KLSSESK 385

RESULT 23
AAG36663
ID AAG36663 standard; Protein: 478 AA.
XX
AC AAG36663;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 44963.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137282.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 03-AUG-1999; 99US-0146388.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156566.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

Best Local Similarity 5.2%; Score 7; DB 21; Length 478;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 64 KLSSEK 70
Db 391 KLSSEK 397

RESULT 24

AAG39611
ID AAG39611 standard; Protein: 478 AA.

XX AAG39611;

XX 18-OCT-2000 (first entry)

```
XX Arabidopsis thaliana protein fragment SEQ ID NO: 49037.
DE
XX Protein identification: signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX Arabidopsis thaliana.
OS
XX EP103405-A2.
PN
XX 06-SEP-2000.
PD
XX
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132483.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144352.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148655.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
```

```

Query Match          5.2%; Score 7; DB 21; Length 478;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY      64 KLSSESK 70
        |||||
Db       391 KLSSESK 397

RESULT 25
AAG36662
ID AAG36662 standard; Protein: 614 AA.
XX
AC AAG36662;
XX
AMG36662;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 44962.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX

```

OS	XX	Arabidopsis thaliana.	
XX	XX	EP103405-A2.	
XX	XX	06-SEP-2000.	
XX	XX	25-FEB-2000;	2000EP-0301439
XX	XX	25-FEB-1999;	99US-0121825
PR	PR	05-MAR-1999;	99US-0123180
PR	PR	09-MAR-1999;	99US-0123348
PR	PR	23-MAR-1999;	99US-0125788
PR	PR	25-MAR-1999;	99US-0126264
PR	PR	01-APR-1999;	99US-0126785
PR	PR	01-APR-1999;	99US-0127462
PR	PR	08-APR-1999;	99US-0128234
PR	PR	16-APR-1999;	99US-0128714
PR	PR	19-APR-1999;	99US-0129845
PR	PR	21-APR-1999;	99US-0130077
PR	PR	23-APR-1999;	99US-0130449
PR	PR	23-APR-1999;	99US-0130510
PR	PR	28-APR-1999;	99US-0130891
PR	PR	30-APR-1999;	99US-0131449
PR	PR	03-MAY-1999;	99US-0132048
PR	PR	04-MAY-1999;	99US-0132407
PR	PR	05-MAY-1999;	99US-0133484
PR	PR	06-MAY-1999;	99US-0133485
PR	PR	07-MAY-1999;	99US-0133487
PR	PR	11-MAY-1999;	99US-0134283
PR	PR	11-MAY-1999;	99US-0134256
PR	PR	14-MAY-1999;	99US-0134218
PR	PR	14-MAY-1999;	99US-0134219
PR	PR	14-MAY-1999;	99US-0134322
PR	PR	18-MAY-1999;	99US-0134370
PR	PR	18-MAY-1999;	99US-0134768
PR	PR	19-MAY-1999;	99US-0134941
PR	PR	20-MAY-1999;	99US-0135124
PR	PR	21-MAY-1999;	99US-0135353
PR	PR	24-MAY-1999;	99US-0135629
PR	PR	25-MAY-1999;	99US-0136021
PR	PR	27-MAY-1999;	99US-0136332
PR	PR	28-MAY-1999;	99US-0136782
PR	PR	01-JUN-1999;	99US-0137222
PR	PR	03-JUN-1999;	99US-0137528
PR	PR	04-JUN-1999;	99US-0137502
PR	PR	07-JUN-1999;	99US-0137724
PR	PR	10-JUN-1999;	99US-0138094
PR	PR	10-JUN-1999;	99US-0138187
PR	PR	10-JUN-1999;	99US-0139111
PR	PR	16-JUN-1999;	99US-0139452
PR	PR	16-JUN-1999;	99US-0139453
PR	PR	17-JUN-1999;	99US-0139492
PR	PR	18-JUN-1999;	99US-0139454
PR	PR	18-JUN-1999;	99US-0139455
PR	PR	18-JUN-1999;	99US-0139456
PR	PR	18-JUN-1999;	99US-0139457
PR	PR	18-JUN-1999;	99US-0139458
PR	PR	18-JUN-1999;	99US-0139459
PR	PR	18-JUN-1999;	99US-0139460
PR	PR	18-JUN-1999;	99US-0139461
PR	PR	18-JUN-1999;	99US-0139462
PR	PR	18-JUN-1999;	99US-0139463
PR	PR	18-JUN-1999;	99US-0139750
PR	PR	21-JUN-1999;	99US-0139817
PR	PR	22-JUN-1999;	99US-0139817
PR	PR	23-JUN-1999;	99US-0140353
PR	PR	23-JUN-1999;	99US-0140354
PR	PR	28-JUN-1999;	99US-0140695
PR	PR	29-JUN-1999;	99US-0140823
PR	PR		99US-0140891

PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143352.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

Query Match 5.2%; Score 7; DB 21; Length 614;
Best Local Similarity 100.0%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

QY 64 KLSSESK 70
|||||||
Db 527 KLSSESK 533

RESULT 26
AAG39610
ID AAG39610 standard; Protein; 618 AA.
XX
AC AAG39610;
XX
DT 18-OCT-2000 (First entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49036.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.

XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132487.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0136229.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139454.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.

```

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match      5.2%; Score 7; DB 21; Length 618;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 64 KLSSESK 70
   |||||
Db 531 KLSSESK 537

```

RESULT 27

```

ABB93742
ID ABB93742 standard; Protein; 618 AA.

```

```

XX ABB93742;

```

```

XX 31-MAY-2002 (first entry)

```

```

DE Herbicidally active polypeptide SEQ ID NO 2953.

```

```

XX Herbicidal; plant; agriculture; herbicide.

```

```

XX Arabidopsis thaliana.

```

```

XX WO200210210-A2.

```

```

XX 07-FEB-2002.

```

```

XX 28-AUG-2001; 2001WO-EP09892.

```

```

XX 28-AUG-2001; 2001WO-EP09892.

```

```

XX (FARB ) BAYER AG.

```

```

XX Tietjen K, Weidner M;

```

```

XX WPI; 2002-269010/31.

```

```

XX

```

```

PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 2953; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
SQ Sequence 618 AA;

```

```

Query Match      5.2%; Score 7; DB 23; Length 618;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 64 KLSSESK 70
   |||||
Db 531 KLSSESK 537

```

RESULT 28

```

ABB48009
ID ABB48009 standard; Protein; 928 AA.

```

```

XX ABB48009;

```

```

XX 05-FEB-2002 (first entry)

```

```

DE Listeria monocytogenes protein #713.

```

```

XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
XX vitamin B12; bacterial infection; disease.

```

```

XX Listeria monocytogenes.

```

```

XX WO200177335-A2.

```

```

XX 18-OCT-2001.

```

```

XX 11-APR-2001; 2001WO-FR01118.

```

```

XX 11-APR-2000; 2000FR-0004629.

```

```

XX (INST ) INST PASTEUR.

```

```

XX Buchrieser C, Frangeul L, Rusanlok C, Eshti H, Dehoux P,
XX Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P,
XX Daniels J, Goebel W, Krefelt J, Kuhn M, Ng E, Vazquez-Boland JA;
XX Dominguez-Bernal G, Garrido-Garcia P, Tietz-Martinez A, Amend A;
XX Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
XX Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
XX Madueno E, De Pablos B, Wehlant J, Kaerst U, Entian K, Hauf J;
XX Rose M, Voss H;
XX
XX

```

```

XX WPI; 2002-010914/01.

```

```

XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
XX and prevention of Listeria and related bacterial infections, and
XX related polypeptides -
XX
XX

```

```

XX Claim 6; SEQ ID NO 714; 192pp; French.

```

```

XX The present invention relates to the genome sequence of Listeria
XX monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
XX it are useful for selecting probes and primers for detecting genes in L.
XX

```

CC monocytes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by L.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 CC Sequence 928 AA:

Query Match 5.2%; Score 7; DB 23; Length 928;
 Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 IAAAKLS 66
 |||||
 Db 520 IAAAKLS 526

RESULT 29

AAW69740
 ID AAW69740 standard; protein: 992 AA.

XX AC AAW69740;

DT 26-OCT-1998 (first entry)

XX DE SAPAP1 protein.

KW Human; SAPAP1; SAPAP2; animal protein; PSD-95/SAP90; diagnosis;
 KW nervous disease; functional interference; structural interference;
 KW membrane associated guanylate kinase; neuronal disease.
 OS Homo sapiens.

PN JP10201477-A.

PD 04-AUG-1998.

PF 24-JAN-1997; 97JP-0011714.

PR 24-JAN-1997; 97JP-0011714.

PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
 PA (TAKE/) TAKEUCHI M.

DR WPI: 1998-474491/41.

PT New protein SAPAP1 - used for, e.g. diagnosis and prevention of
 PT various neuronal diseases

PS Claim 1; Page 4-7; 12pp; Japanese.

CC The present sequence represents the SAPAP1 protein having a 992 amino
 CC acid (aa) sequence. Also described in the present invention are: (1) an
 CC animal protein having an aa sequence substantially homologous to SAPAP1;
 CC (2) cDNA sequence encoding SAPAP1, or an aa sequence substantially
 CC homologous to SAPAP1, and (3) a genomic DNA sequence hybridised to the
 CC cDNA or its partial sequence. SAPAP1 is a novel animal protein specific
 CC for PSD-95/SAP90 and its related protein, and may be useful for the
 CC diagnosis, prevention and treatment of various neuronal diseases caused
 CC by functional or structural interference of nervous system.
 CC
 CC Sequence 992 AA:

Query Match 5.2%; Score 7; DB 19; Length 992;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 SLTELT 96
 |||||
 Db 389 SLTELT 395

RESULT 30

AAW69742
 ID AAW69742 standard; protein: 992 AA.

XX AC AAW69742;

DT 26-OCT-1998 (first entry)

XX DE SAPAP2 protein.

KW Human; SAPAP2; SAPAP1; animal protein; PSD-95/SAP90; diagnosis;
 KW nervous disease; functional interference; structural interference;
 KW membrane associated guanylate kinase; neuronal disease.
 OS Homo sapiens.

PN JP10201478-A.

PD 04-AUG-1998.

PF 24-JAN-1997; 97JP-0011715.

PR 24-JAN-1997; 97JP-0011715.

PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
 PA (TAKE/) TAKEUCHI M.

DR WPI: 1998-474492/41.

PT DNA encoding new animal protein SAPAP 2 - useful for diagnosis and
 PT treatment of nervous system diseases

PS Claim 1; Page 7-9; 11pp; Japanese.

CC The present sequence represents the SAPAP2 protein. Also described in
 CC the present invention is: (A) an animal protein having an amino acid
 CC sequence substantially the same as SAPAP2; (B) a cDNA sequence encoding
 CC the amino acid sequence of SAPAP2 or (C) a genomic DNA sequence
 CC substantially same as SAPAP2; and (D) a genomic DNA sequence hybridised
 CC by the above cDNA or its partial sequence. SAPAP2 is a new animal
 CC protein which combines specifically with PSD-95/SAP90 and its related
 CC protein and is useful for the diagnosis, prevention and treatment of
 CC various nervous diseases caused by functional or structural interference
 CC of nervous system.
 CC
 CC Sequence 992 AA:

SQ Sequence 992 AA:

Query Match 5.2%; Score 7; DB 19; Length 992;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 SLTELT 96
 |||||
 Db 389 SLTELT 395

RESULT 31

ABB68996
 ID ABB68996 standard; protein: 1848 AA.

XX AC ABB68996;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 33780.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 PI
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL13099.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 33780; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-ABL30511), expressed DNA
 CC sequences (AB101840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1848 AA;
 XX
 Query Match 5.2%; Score 7; DB 22; Length 1848;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 31 SSANLFT 37
 DB 1722 SSANLFT 1728
 XX
 RESULT 32
 ABB58665
 ID ABB58665 standard; Protein; 4796 AA.
 XX
 AC ABB58665;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 2787.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX

PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 PI
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL02768.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 2787; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-ABL30511), expressed DNA
 CC sequences (AB101840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 4796 AA;
 XX
 Query Match 5.2%; Score 7; DB 22; Length 4796;
 Best Local Similarity 100.0%; Pred. No. 9.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 101 PNARTEV 107
 DB 2208 PNARTEV 2214
 XX
 RESULT 33
 AAR94612
 ID AAR94612 standard; peptide; 11 AA.
 XX
 AC AAR94612;
 XX
 DT 21-OCT-1996 (first entry)
 XX
 DE Protein kinase binding region on c-Fos.
 XX
 KW Protein kinase; c-Fos; activation; phosphorylation; cell proliferation;
 KW disorder; inhibition.
 XX
 OS Synthetic.
 OS
 PN WO9609835-A1.
 PN
 PD 04-APR-1996.
 XX
 PF 13-FEB-1995; 95WO-US01770.
 XX
 PR 29-SEP-1994; 94US-0315067.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Deng T, Karin M;
 PI
 DR WPI; 1996-200718/20.
 XX
 PT c-Fos regulating protein kinase phosphorylates c-Fos - acts to
 PT potentiates its activity, used to treat a cell proliferation
 PT disorder
 XX
 PS Claim 23; Page 39; 55pp; English.

XX This sequence comprises the protein binding region on c-Fos which
CC corresponds to amino acids 226-236 of the mature protein. The
CC peptide can be useful for the competitive inhibition of naturally
CC occurring c-Fos in situations where it may be desirable to decrease
CC the amount of c-Fos activation by phosphorylation by a protein
CC kinase.
XX

SO Sequence 11 AA;
Query Match 4.5%; Score 6; DB 17; Length 11;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 21 ASTPES 26
|||||
Db 5 ASTPES 10

RESULT 34
AAU12145
ID AAU12145 standard; peptide: 11 AA.
XX
AC AAU12145;
XX
DT 26-MAR-2002 (first entry)
XX
DE Arabidopsis RPW8.2 conserved peptide motif #4.
XX
KM RPW8.2: resistance to powdery mildew; Erysiphe cichoracearum;
KW E. cruciferarum; E. orontii; Oidium lycopersici; transgenic plant.
XX Arabidopsis thaliana.
OS
XX
PN MO200198479-A2.
XX
PD 27-DEC-2001.
XX
PF 19-JUN-2001; 2001WO-GB02693.
XX
PR 20-JUN-2000; 2000GB-0015122.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Xiao SY, Turner JG, Coleman M, Ellwood S;
XX
DR WPI: 2002-114581/15.
XX
XX Plant resistance genes from Arabidopsis thaliana and related homologues
XX useful for manipulating the resistance of plants to powdery mildew -
XX
XX Claim 7; Page 94; 104pp; English.

The invention relates to a nucleic acid molecule comprising an RPW
cichoracearum, E. cruciferarum, E. orontii and Oidium lycopersici)
nucleotide sequence encoding an RPW resistance polypeptide having an
N-terminal transmembrane domain and a coiled coil domain and which is
capable of recognising and activating in a plant into which the nucleic
acid is introduced a specific defence response to challenge with a
powdery mildew pathogen. Also included are primers and probes
designed from conserved sequences from the CDNAS for RPW8.1 or
RPW8.2, which may be used to amplify RPW8.1/8.2 sequences, a
recombinant vector comprising the RPW polynucleotide, a host cell
comprising the vector and a transgenic plant comprising the
polynucleotide. The RPW polynucleotide is useful for influencing the
affecting, in particular increasing the degree of resistance of a plant
to a powdery mildew pathogens such as Erysiphe cichoracearum, E.
cruciferarum, E. orontii and Oidium lycopersici. The polynucleotide,
primers or complement of the polynucleotide are useful for identifying,
cloning or determining the presence of the polynucleotide sequence within a
plant. The vector is useful for transforming a host cell for producing a
transgenic plant such as wheat, barley, tomato or Nicotiana sp., by

CC introducing the vector into a host cell and optionally causing or
CC allowing recombination between the vector and the host cell genome so as
CC to transform the host cell. The present sequence is a RPW8.2
CC conserved peptide motif used to design degenerate PCR primers.
XX

SO Sequence 11 AA;
Query Match 4.5%; Score 6; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 70 KAKISE 75
|||||
Db 6 KAKISE 11

RESULT 35
AAV36488
ID AAV36488 standard; Protein; 26 AA.
XX
AC AAV36488;
XX
DT 17-SEP-1999 (first entry)
XX
DE Fragment of human secreted protein encoded by gene 22.
XX
KW Human; secreted protein; cancer; tumour; developmental abnormality;
KW foetal deficiency; blood disorder; immune system disorder; inflammation;
KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
KW digestive disorder; endocrine disorder; infection; AIDS.
XX
OS Homo sapiens.
XX
XX WO9931117-A1.
XX
PD 24-JUN-1999.
XX
PF 17-DEC-1998; 98WO-US27059.
XX
PR 19-DEC-1997; 97US-0068369.
PR 18-DEC-1997; 97US-0068006.
PR 18-DEC-1997; 97US-0068007.
PR 18-DEC-1997; 97US-0068008.
PR 18-DEC-1997; 97US-0068053.
PR 18-DEC-1997; 97US-0068054.
PR 18-DEC-1997; 97US-0068057.
PR 18-DEC-1997; 97US-0068064.
PR 18-DEC-1997; 97US-0070923.
PR 19-DEC-1997; 97US-0068169.
PR 19-DEC-1997; 97US-0068365.
PR 19-DEC-1997; 97US-0068367.
PR 19-DEC-1997; 97US-0068368.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Carter KC, Duan RD, Feng P, Ferrie AM, Florence C;
PI Florence K, Greene JM, Janat F, Kyaw H, Moore PA;
PI Ni J, Rosen CA, Ruben SM, Shi Y, Soppet DR, Wei Y;
PI Yu G;
XX
XX WPI: 1999-418749/35.
XX
XX New isolated human genes encoding secreted polypeptides
XX
XX Disclosure: Page 463; 537pp; English.

AAV97916 to AAV98029 represent 110 isolated human secreted protein
genes. AAV36224 to AAV36727 represent the secreted proteins encoded by
the 110 human genes. The genes and their corresponding secreted
polypeptides are useful for preventing, treating or ameliorating medical
conditions, e.g. by protein or gene therapy. Also pathological conditions

CC can be diagnosed by determining the amount of the new polypeptides in a
 CC sample or by determining the presence of mutations in the new genes.
 CC Specific uses are described for each of the 110 genes, based on which
 CC tissues they are most highly expressed in, and include developing
 CC products for the diagnosis or treatment of cancer, tumours, developmental
 CC abnormalities and foetal deficiencies, blood disorders, diseases of the
 CC immune system, autoimmune diseases, inflammation, allergies, Alzheimer's
 CC and cognitive disorders, schizophrenia, arthritis, asthma, psoriasis,
 CC sepsis, skin disorders, atherosclerosis, diabetes, cardiovascular
 CC disorders, kidney disorders, digestive/endocrine disorders, infections
 CC and AIDS. The polypeptides are also useful for identifying their binding
 CC partners. The sequences given in AAX97907 to AAX97915 and AAY36223 are
 CC used in the exemplification of the present invention.

CC Sequence 26 AA:

Query Match 4.5%; Score 6; DB 20; Length 26;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 115 SLKPCM 120
 |||||
 Db 11 SLKPCM 16

RESULT 36

ABB32196
 ID ABB32196 standard; Peptide; 29 AA.

AC ABB32196;

DT 01-FEB-2002 (first entry)

DE Peptide #4847 encoded by breast cell single exon nucleic acid probe.

XX Human; microarray; single exon probe; gene expression; breast;
 KM disease; cancer.

OS Homo sapiens.

PN WO200157271-A2.

XX 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US00662.

PF 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-496933/54.

DR New spatially-addressable set of single exon nucleic acid probes,
 PT useful for measuring gene expression in sample derived from human
 PT breast, comprises number of single exon nucleic acid probes -
 XX

PS Claim 27; SEQ ID NO 15164; 327pp + sequence listing; English.

XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting
 CC the probes with a collection of detectably labelled nucleic acids
 CC derived from mRNA of human breast, and then measuring the label
 CC bound to each probe of the microarray. The probes are useful for
 CC verifying the expression of regions of genomic DNA predicted to

CC encode proteins. They are useful for gene discovery, and for
 CC determining predisposition and/or prognosing breast disease. Gene
 CC expression analysis is useful for assessing the toxicity of chemical
 CC agents on cells. The microarray of this invention presents a far greater
 CC diversity of probes for measuring gene expression, with far less bias
 CC than expressed sequence tag microarrays. The method is suitable for
 CC rapid production of functional information from genomic sequence. The
 CC present sequence is a peptide encoded by a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published-pct-sequences.

CC Sequence 29 AA:

Query Match 4.5%; Score 6; DB 22; Length 29;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 109 OKIVRH 114
 |||||
 Db 5 OKIVRH 10

RESULT 37

ABG40230
 ID ABG40230 standard; Peptide; 29 AA.

AC ABG40230;

DT 19-AUG-2002 (first entry)

DE Human peptide encoded by genome-derived single exon probe SEQ ID 29895.

XX Human; single exon probe; asthma; lung cancer; COPD; ID;
 KM chronic obstructive pulmonary disease; interstitial lung disease;
 KM familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KM tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KM Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KM pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KM primary ciliary dyskinesia; pulmonary hypertension;
 KM hyaline membrane disease.

XX Homo sapiens.

OS WO200186003-A2.

PN 15-NOV-2001.

PD 30-JAN-2001; 2001WO-US00665.

PF 04-FEB-2000; 2000US-180312P.

PR 26-MAY-2000; 2000US-207456P.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-234687P.

PR 27-SEP-2000; 2000US-236359P.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2002-114183/15.

DR Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 PT

PS Claim 27; SEQ ID NO 29895; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived

CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe.
 CC In the above mentioned microarray: assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 29 AA;

Query Match
 Best Local Similarity 4.5%; Score 6; DB 23; Length 29;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 109 QKIVRH 114
 |||||
 Db 5 QKIVRH 10

RESULT 38
 AAB32011
 ID AAB32011 standard; peptide: 34 AA.
 AC AAB32011;
 XX
 DT 14-FEB-2001 (first entry)
 XX
 DE Human secreted protein SEQ ID NO: 69.
 XX
 DE
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotoxic; antidiabetic; antineoplastic; antitumor;
 KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200058350-A1.
 XX
 PD 05-OCT-2000.
 XX
 PF 22-MAR-2000; 2000MO-US07483.

XX
 PR 26-MAR-1999; 99US-0126596.
 PR 22-DEC-1999; 99US-0171552.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Komatsoulis G;
 XX
 DR WPI: 2000-602357/57.
 DR N-PSDB; AAC66419.
 XX
 PT Nucleic acid molecules encoding human secreted proteins, used in
 PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
 PT Parkinson's diseases and cancers -
 PS
 XX
 PS Claim 11; Page 368; 423pp; English.

CC Sequences AAB32002-B32050 represent the amino acid sequences of 49
 CC human secreted proteins encoded by the genes AAC66410-C66458. The genes
 CC and proteins are useful for preventing, ameliorating or treating medical
 CC conditions, e.g. by protein or gene therapy. The genes are isolated from
 CC a range of human tissues disclosed in the specification. The nucleic
 CC acids, proteins, antibodies and (ant)agonists are useful in the
 CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer, and other cancers of the adrenal gland; bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC Crohn's disease, autoimmune thyroiditis, diabetes mellitus,
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC
 SQ Sequence 34 AA;

Query Match
 Best Local Similarity 4.5%; Score 6; DB 21; Length 34;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 94 LTTVAI 99
 |||||
 Db 24 LTTVAI 29

RESULT 39
 AAG80045
 ID AAG80045 standard; peptide: 39 AA.
 AC AAG80045;
 XX
 DT 17-JAN-2002 (first entry)
 XX
 DE Chemokine peptide fragment CXCR1.
 XX
 DE
 XX
 KW Chemokine; tumour diagnosis; colorectal; prostatic; organ rejection;
 KW inflammation; autoimmune disease; metastasis; bronchial asthma; lupus;
 KW chronic bowel inflammation; rheumatoid arthritis; cytostatic;
 KW antineoplastic; antitumour; immunosuppressive; dermatological;
 KW antirheumatic; antiallergic.
 XX
 OS Homo sapiens.
 XX
 PN WO200172830-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 02-APR-2001; 2001WO-EP03708.
 XX
 PR 31-MAR-2000; 2000DE-1016013.
 XX
 PA (TFP-) IPF PHARM GMBH.
 PA (FORS/) FORSMANN U.

XX Forssmann W, Adermann K, Heiland A, Spodsberg N;
 PI WPI: 2001-626256/72.
 DR
 XX
 PT Diagnostic agent containing two or more receptor-specific ligands,
 PT useful for detecting tumors, inflammation etc., also therapeutic use of
 PT ligand inhibitors
 PS
 XX Claim 24, Page 7; 26pp; German.
 CC This invention describes a novel diagnostic agent (A) comprising at least
 CC two different ligands (I) for receptors (II) that are implicated in
 CC disease. (A) are used for the diagnosis of tumors (especially colorectal
 CC or prostatic), organ rejection, inflammation and autoimmune diseases.
 CC Also inhibitors of (I) are used therapeutically against tumors (and their
 CC metastases), inflammation (particularly bronchial asthma or chronic bowel
 CC inflammation), or autoimmune diseases (rheumatoid arthritis or lupus),
 CC where the (cardio)vascular, lymphatic, respiratory, nervous, digestive,
 CC endocrine, motor or urogenital systems or skin are affected, and bone
 CC marrow diseases. The products of the invention are chemokine derivatives
 CC which have cytostatic, antiinflammatory, antisthmatic,
 CC immunosuppressive, dermatological, antirheumatic, antiarthritic.
 CC Chemokines act on specific tumor and inflammatory cells through a
 CC constellation of chemokine receptors (CR), which control migration and
 CC proliferation of these cells. AAG8045-AAG80128 represent human chemokine
 CC fragments used to illustrate the method of the invention.
 CC
 SO Sequence 39 AA:

Query Match 4.5%; Score 6; DB 22; Length 39;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 118 PCMLET 123
 |||||
 DB 29 PCMLET 34

RESULT 40

ABB35095
 ID ABB35095 standard; Peptide: 42 AA.

XX
 AC ABB35095;

DT 04-FEB-2002 (first entry)

XX Peptide #2601 encoded by human foetal liver single exon probe.

DE Human; foetal liver; gene expression; single exon nucleic acid probe.

XX Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00669.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI: 2001-483447/52.

PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 XX
 XX Claim 27, SEQ ID NO 27730; 639pp + sequence listing; English.
 PS
 XX
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at http://wipo.int/pub/published_pct_sequences.
 CC
 SO Sequence 42 AA:

Query Match 4.5%; Score 6; DB 22; Length 42;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLPTS 38
 |||||
 DB 14 ANLPTS 19

Search completed: April 28, 2003, 16:13:03
 Job time : 73 secs

This Page Blank (uspto)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 16:11:58 ; Search time is Seconds
(without alignments)
262.845 Million cell updates/sec

Title: US-10-048-196-2

Perfect score: 134

Sequence: 1 MKKILYVATLMTAFITLASC.....SLKPCMLETVNATIVPTTR 134

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

Issued Patents.AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5.2	40	4	US-09-082-279B-1138	Sequence 1138, Ap
2	7	40	4	US-09-315-304B-1138	Sequence 1138, Ap
3	4.5	11	1	US-08-707-873-1	Sequence 1, Appli
4	4.5	11	2	US-08-707-874-1	Sequence 1, Appli
5	4.5	11	2	US-08-315-067-1	Sequence 1, Appli
6	4.5	11	3	US-09-193-797-1	Sequence 1, Appli
7	4.5	11	5	PCT-US95-01770-1	Sequence 1, Appli
8	4.5	39	4	US-08-487-795A-24	Sequence 24, Appl
9	4.5	39	4	US-08-121-105B-24	Sequence 24, Appl
10	4.5	39	5	PCT-US94-10356-27	Sequence 27, Appl
11	4.5	76	2	US-08-469-537A-38	Sequence 38, Appl
12	4.5	93	6	5169835-26	Patent No. 5169835
13	4.5	102	2	US-08-750-856A-16	Sequence 16, Appl
14	4.5	120	4	US-09-107-858-23	Sequence 23, Appl
15	4.5	144	6	5169835-8	Patent No. 5169835
16	4.5	154	4	US-09-228-986-89	Sequence 89, Appl
17	4.5	160	4	US-07-847-010-3	Sequence 3, Appli
18	4.5	161	4	US-09-615-192A-388	Sequence 388, App
19	4.5	195	4	US-09-370-838-125	Sequence 125, App
20	4.5	230	6	5169835-13	Patent No. 5169835
21	4.5	244	1	US-08-318-947A-21	Sequence 21, Appl
22	4.5	244	2	US-08-795-303-21	Sequence 21, Appl
23	4.5	259	2	US-07-857-224B-49	Sequence 49, Appl
24	4.5	271	2	US-08-467-265-2	Sequence 2, Appli
25	4.5	271	4	US-08-467-265-2	Sequence 2, Appli
26	4.5	271	4	US-09-407-891-2	Sequence 2, Appli
27	4.5	273	2	US-08-701-191A-32	Sequence 32, Appl

28	4.5	277	4	US-09-090-793-3	Sequence 3, Appli
29	4.5	296	1	US-08-241-465B-21	Sequence 21, Appl
30	4.5	297	4	US-09-173-300-26	Sequence 26, Appl
31	4.5	300	2	US-08-946-528-7	Sequence 7, Appli
32	4.5	307	4	US-09-173-300-15	Sequence 15, Appl
33	4.5	315	3	US-09-135-639-4	Sequence 19, Appl
34	4.5	334	1	US-08-241-465B-19	Sequence 19, Appl
35	4.5	334	1	US-08-241-465B-20	Sequence 20, Appl
36	4.5	345	2	US-08-758-621-14	Sequence 14, Appl
37	4.5	345	4	US-09-107-858-14	Sequence 14, Appl
38	4.5	348	4	US-09-134-001C-5513	Sequence 513, Ap
39	4.5	350	1	US-08-202-056-1	Sequence 1, Appli
40	4.5	350	1	US-08-076-093A-2	Sequence 2, Appli
41	4.5	350	1	US-08-450-393A-7	Sequence 7, Appli
42	4.5	350	1	US-08-410-453A-1	Sequence 1, Appli
43	4.5	350	1	US-08-701-265-2	Sequence 2, Appli
44	4.5	350	1	US-08-410-454A-1	Sequence 1, Appli
45	4.5	350	2	US-08-284-586-2	Sequence 2, Appli
46	4.5	350	2	US-08-410-456A-1	Sequence 1, Appli
47	4.5	350	2	US-08-805-478-2	Sequence 2, Appli
48	4.5	350	2	US-08-802-627A-2	Sequence 2, Appli
49	4.5	350	2	US-08-801-238-2	Sequence 2, Appli
50	4.5	350	2	US-08-801-228-2	Sequence 2, Appli
51	4.5	350	3	US-09-104-296-2	Sequence 2, Appli
52	4.5	350	4	US-08-446-669-7	Sequence 7, Appli
53	4.5	350	5	PCT-US95-00476-7	Sequence 7, Appli
54	4.5	353	1	US-08-118-270-45	Sequence 45, Appl
55	4.5	353	2	PCT-US93-08528-45	Sequence 45, Appl
56	4.5	364	2	US-08-444-646-3	Sequence 3, Appli
57	4.5	366	3	US-08-945-056-6	Sequence 6, Appli
58	4.5	410	4	US-08-725-758A-4	Sequence 4, Appli
59	4.5	419	6	5169835-2	Patent No. 5169835
60	4.5	426	4	US-08-725-758A-2	Sequence 2, Appli
61	4.5	431	4	US-08-845-258-34	Sequence 34, Appl
62	4.5	431	4	US-08-900-571-34	Sequence 34, Appl
63	4.5	431	4	US-08-723-142A-34	Sequence 34, Appl
64	4.5	431	4	US-09-528-784A-34	Sequence 34, Appl
65	4.5	436	3	US-08-486-099A-94	Sequence 94, Appl
66	4.5	436	3	US-08-360-107A-104	Sequence 104, App
67	4.5	436	3	US-08-484-223B-94	Sequence 94, Appl
68	4.5	436	3	US-08-919-537-94	Sequence 94, Appl
69	4.5	436	3	US-08-475-668A-94	Sequence 94, Appl
70	4.5	436	3	US-08-485-551A-94	Sequence 94, Appl
71	4.5	436	3	US-08-471-913A-94	Sequence 94, Appl
72	4.5	436	4	US-08-485-264A-94	Sequence 94, Appl
73	4.5	436	4	US-08-584-760A-67	Sequence 67, Appl
74	4.5	436	4	US-08-474-349A-94	Sequence 94, Appl
75	4.5	436	4	US-08-255-208A-30	Sequence 30, Appl
76	4.5	440	4	US-08-584-760A-1	Sequence 1, Appli
77	4.5	451	4	US-09-134-001C-4461	Sequence 4461, Ap
78	4.5	487	1	US-08-249-112-4	Sequence 4, Appli
79	4.5	487	5	PCT-US95-06556-4	Sequence 4, Appli
80	4.5	497	4	US-09-377-557-10	Sequence 10, Appl
81	4.5	501	1	US-08-722-001-14	Sequence 14, Appl
82	4.5	501	2	US-08-467-568-9	Sequence 9, Appli
83	4.5	501	2	US-09-030-582-9	Sequence 9, Appli
84	4.5	501	4	US-09-688-415-7	Sequence 7, Appli
85	4.5	521	2	US-08-406-855A-19	Sequence 19, Appl
86	4.5	521	3	US-09-206-899-19	Sequence 19, Appl
87	4.5	533	1	US-07-820-011A-2	Sequence 2, Appli
88	4.5	533	5	PCT-US93-00445-2	Sequence 2, Appli
89	4.5	536	1	US-07-820-011A-4	Sequence 4, Appli
90	4.5	536	4	US-08-426-509A-13	Sequence 13, Appl
91	4.5	536	5	PCT-US93-00445-4	Sequence 4, Appli
92	4.5	536	5	PCT-US95-00508-13	Sequence 13, Appl
93	4.5	545	2	US-08-467-822-30	Sequence 30, Appl
94	4.5	545	4	US-08-437-697-30	Sequence 30, Appl
95	4.5	545	4	US-08-466-248-30	Sequence 30, Appl
96	4.5	551	2	US-08-700-548-4	Sequence 4, Appli
97	4.5	553	6	5310678-1	Patent No. 5310678
98	4.5	559	2	US-08-406-855A-20	Sequence 20, Appl
99	4.5	559	3	US-09-206-899-20	Sequence 20, Appl
100	4.5	560	4	US-09-688-415-8	Sequence 8, Appli

101	6	4.5	572	1	US-08-334-698-2	Sequence 2, Appl1	174	5	3.7	11	2	US-08-827-009-1	Sequence 1, Appl1
102	6	4.5	572	1	US-08-228-932-2	Sequence 2, Appl1	175	5	3.7	11	3	US-08-974-899-16	Sequence 16, Appl1
103	6	4.5	572	1	US-08-468-939-2	Sequence 2, Appl1	176	5	3.7	11	3	US-08-422-093-3	Sequence 3, Appl1
104	6	4.5	572	1	US-08-722-001-30	Sequence 30, Appl1	177	5	3.7	11	3	US-08-891-845-5	Sequence 5, Appl1
105	6	4.5	572	2	US-08-406-855A-2	Sequence 2, Appl1	178	5	3.7	11	3	US-08-422-112-3	Sequence 3, Appl1
106	6	4.5	572	2	US-08-722-190-2	Sequence 2, Appl1	179	5	3.7	11	3	US-08-822-940-5	Sequence 4, Appl1
107	6	4.5	572	3	US-08-244-354-2	Sequence 2, Appl1	180	5	3.7	11	4	US-08-822-940-5	Sequence 5, Appl1
108	6	4.5	572	3	US-09-206-899-2	Sequence 2, Appl1	181	5	3.7	11	4	US-08-822-940-5	Sequence 4, Appl1
109	6	4.5	572	4	US-09-444-783-2	Sequence 2, Appl1	182	5	3.7	11	4	US-08-822-940-5	Sequence 5, Appl1
110	6	4.5	572	4	US-09-688-415-2	Sequence 2, Appl1	183	5	3.7	11	6	5200320-35	Sequence 27, Appl1
111	6	4.5	582	5	PCR-US95-04203-2	Sequence 2, Appl1	184	5	3.7	13	1	US-08-619-645-5	Patent No. 5200320
112	6	4.5	591	1	US-08-261-086-2	Sequence 2, Appl1	185	5	3.7	13	1	US-08-619-645-6	Sequence 5, Appl1
113	6	4.5	591	3	US-08-432-473-4	Sequence 4, Appl1	186	5	3.7	13	1	US-08-422-101-15	Sequence 6, Appl1
114	6	4.5	648	4	US-09-199-637A-221	Sequence 4, Appl1	187	5	3.7	13	1	US-08-422-091-15	Sequence 15, Appl1
115	6	4.5	667	4	US-09-315-127-5	Sequence 5, Appl1	188	5	3.7	13	2	US-08-634-493-5	Sequence 15, Appl1
116	6	4.5	667	4	US-09-315-127-5	Sequence 5, Appl1	189	5	3.7	13	2	US-08-634-493-6	Sequence 6, Appl1
117	6	4.5	727	1	US-08-424-424B-2	Sequence 6, Appl1	190	5	3.7	13	2	US-08-422-093-15	Sequence 15, Appl1
118	6	4.5	727	1	US-08-424-424B-2	Sequence 6, Appl1	191	5	3.7	13	3	US-08-422-093-15	Sequence 15, Appl1
119	6	4.5	727	1	US-08-424-424B-2	Sequence 6, Appl1	192	5	3.7	13	3	US-08-422-093-15	Sequence 15, Appl1
120	6	4.5	737	1	US-08-188-582-16	Sequence 2, Appl1	193	5	3.7	13	4	US-09-405-745-1	Sequence 1, Appl1
121	6	4.5	737	1	US-08-646-715-16	Sequence 2, Appl1	194	5	3.7	13	4	US-09-405-745-1	Sequence 1, Appl1
122	6	4.5	774	4	US-08-270-076A-11	Sequence 16, Appl1	195	5	3.7	15	1	US-08-101-041A-7	Sequence 5, Appl1
123	6	4.5	806	1	US-08-486-099-103	Sequence 11, Appl1	196	5	3.7	15	1	US-08-218-025A-103	Sequence 7, Appl1
124	6	4.5	856	3	US-08-484-223B-103	Sequence 103, App	197	5	3.7	15	2	US-08-967-101-170	Sequence 103, App
125	6	4.5	856	3	US-08-919-597-103	Sequence 103, App	198	5	3.7	15	2	US-08-967-101-170	Sequence 170, App
126	6	4.5	856	3	US-08-475-668A-103	Sequence 103, App	199	5	3.7	15	3	US-08-592-541-170	Sequence 170, App
127	6	4.5	856	3	US-08-475-668A-103	Sequence 103, App	200	5	3.7	15	4	US-09-124-698-170	Sequence 170, App
128	6	4.5	856	3	US-08-471-913A-103	Sequence 103, App	201	5	3.7	15	4	US-08-496-841C-167	Sequence 167, App
129	6	4.5	856	4	US-08-485-264A-103	Sequence 103, App	202	5	3.7	15	4	US-09-405-745-2	Sequence 2, Appl1
130	6	4.5	856	4	US-08-485-264A-103	Sequence 103, App	203	5	3.7	15	4	US-09-405-745-2	Sequence 2, Appl1
131	6	4.5	856	4	US-08-474-349A-103	Sequence 103, App	204	5	3.7	15	5	US-08-124-698-170	Sequence 45, Appl1
132	6	4.5	857	1	US-08-220-151-10	Sequence 10, App	205	5	3.7	15	5	PCR-US93-11703-45	Sequence 192, App
133	6	4.5	857	1	US-08-413-118-10	Sequence 10, App	206	5	3.7	15	5	PCR-US93-11703-45	Sequence 40, Appl1
134	6	4.5	857	3	US-08-804-439A-18	Sequence 18, Appl1	207	5	3.7	16	1	US-08-218-025A-192	Sequence 40, Appl1
135	6	4.5	857	3	US-08-360-107A-113	Sequence 10, App	208	5	3.7	16	4	US-09-171-705-40	Sequence 3, Appl1
136	6	4.5	857	3	US-08-473-446-10	Sequence 10, App	209	5	3.7	16	4	US-09-171-705-41	Sequence 38, Appl1
137	6	4.5	857	3	US-08-720-229-18	Sequence 18, Appl1	210	5	3.7	20	1	US-08-405-745-3	Sequence 39, Appl1
138	6	4.5	1008	2	US-08-680-326-30	Sequence 30, Appl1	211	5	3.7	20	1	US-08-218-025A-39	Sequence 49, Appl1
139	6	4.5	1013	2	US-08-686-650-5	Sequence 5, Appl1	212	5	3.7	20	1	US-08-306-116A-4	Sequence 1, Appl1
140	6	4.5	1013	2	US-09-021-287-5	Sequence 5, Appl1	213	5	3.7	20	2	US-08-407-252-1	Sequence 13, Appl1
141	6	4.5	1013	3	US-08-921-408-2	Sequence 2, Appl1	214	5	3.7	20	3	US-08-872-094-13	Sequence 19, Appl1
142	6	4.5	1013	4	US-09-240-473-5	Sequence 2, Appl1	215	5	3.7	20	4	US-08-952-089A-19	Sequence 37, Appl1
143	6	4.5	1013	4	US-09-432-473-2	Sequence 32, Appl1	216	5	3.7	20	4	US-08-455-685-37	Patent No. 5258287
144	6	4.5	1094	4	US-09-268-347-32	Sequence 1, Appl1	217	5	3.7	21	6	5258287-6	Sequence 35, Appl1
145	6	4.5	1388	2	US-08-685-576-1	Sequence 1, Appl1	218	5	3.7	21	3	US-08-356-747C-35	Sequence 35, Appl1
146	6	4.5	1861	2	US-08-790-912-4	Sequence 4, Appl1	219	5	3.7	21	4	US-08-834-130A-33	Sequence 97, Appl1
147	6	4.5	2183	3	US-08-746-111-5	Sequence 5, Appl1	220	5	3.7	22	2	US-08-484-631-97	Sequence 97, Appl1
148	6	4.5	3418	2	US-08-639-501-2	Sequence 2, Appl1	221	5	3.7	22	2	US-08-827-570-97	Sequence 97, Appl1
149	6	4.5	3418	2	US-08-603-753D-4	Sequence 4, Appl1	222	5	3.7	22	2	US-08-827-570-97	Sequence 80, Appl1
150	6	4.5	3418	3	US-09-044-946-2	Sequence 4, Appl1	223	5	3.7	24	1	US-08-118-193-80	Sequence 100, App
151	6	4.5	3418	3	US-09-044-946-2	Sequence 4, Appl1	224	5	3.7	24	1	US-08-484-631-100	Sequence 100, App
152	6	4.5	3418	3	US-08-755-587-44	Sequence 2, Appl1	225	5	3.7	24	2	US-08-827-570-100	Sequence 100, App
153	6	4.5	3418	3	US-09-044-908-2	Sequence 4, Appl1	226	5	3.7	24	2	US-08-827-570-100	Sequence 100, App
154	6	4.5	3418	4	US-09-099-753-4	Sequence 4, Appl1	227	5	3.7	24	4	US-08-952-089A-19	Sequence 69, Appl1
155	6	4.5	3418	4	US-08-986-106-4	Sequence 3159, App	228	5	3.7	24	4	US-08-834-130A-69	Sequence 69, Appl1
156	6	4.5	10182	4	US-09-134-001C-3159	Sequence 31, Appl1	229	5	3.7	24	4	US-08-900-230-34	Sequence 34, Appl1
157	5	3.7	7	1	US-08-288-728-31	Sequence 3, Appl1	230	5	3.7	25	3	US-08-620-151-52	Sequence 52, Appl1
158	5	3.7	8	1	US-08-594-447-3	Sequence 2, Appl1	231	5	3.7	25	4	US-08-927-597-55	Sequence 2, Appl1
159	5	3.7	8	1	US-08-541-964-2	Sequence 17, Appl1	232	5	3.7	25	4	US-08-612-973-55	Sequence 2, Appl1
160	5	3.7	8	4	US-08-665-647-17	Sequence 3, Appl1	233	5	3.7	25	4	US-08-927-597-55	Sequence 55, Appl1
161	5	3.7	9	1	US-08-953-033-3	Sequence 2, Appl1	234	5	3.7	26	2	US-08-493-233-8	Sequence 8, Appl1
162	5	3.7	9	1	US-08-619-645-1	Sequence 1, Appl1	235	5	3.7	26	2	US-08-620-151-52	Sequence 52, Appl1
163	5	3.7	9	2	US-08-619-645-2	Sequence 1, Appl1	236	5	3.7	28	1	US-07-633-964-15	Sequence 15, Appl1
164	5	3.7	9	2	US-08-634-493-1	Sequence 136, App	237	5	3.7	28	1	US-08-386-956-15	Sequence 46, Appl1
165	5	3.7	9	3	US-08-634-493-2	Sequence 239, App	238	5	3.7	28	1	US-09-108-709-46	Sequence 13, Appl1
166	5	3.7	9	3	US-08-159-339A-136	Sequence 249, App	239	5	3.7	30	1	US-09-605-858-13	Sequence 16, Appl1
167	5	3.7	9	3	US-08-159-339A-232	Sequence 3, Appl1	240	5	3.7	30	1	US-07-633-964-16	Sequence 16, Appl1
168	5	3.7	9	3	US-08-159-339A-249	Sequence 3, Appl1	241	5	3.7	30	1	US-08-386-956-16	Sequence 16, Appl1
169	5	3.7	11	1	US-07-977-672-5	Sequence 3, Appl1	242	5	3.7	30	2	US-08-846-762-66	Sequence 66, Appl1
170	5	3.7	11	1	US-08-422-101-3	Sequence 3, Appl1	243	5	3.7	30	3	US-09-108-709-47	Sequence 47, Appl1
171	5	3.7	11	1	US-08-422-091-3	Sequence 3, Appl1	244	5	3.7	31	4	US-09-288-143-149	Sequence 149, App
172	5	3.7	11	1	US-08-422-092-3	Sequence 8, Appl1	245	5	3.7	31	4	US-08-602-999A-106	Sequence 106, App
173	5	3.7	11	2	US-08-788-800-8	Sequence 8, Appl1	246	5	3.7	31	4	US-08-278-665-106	Sequence 106, App

247	5	3.7	31	4	US-09-500-124-106	Sequence 106, App	320	5	3.7	100	1	US-08-422-091-10	Sequence 10, App
248	5	3.7	33	2	US-08-637-759B-312	Sequence 312, App	321	5	3.7	100	2	US-08-422-092-10	Sequence 10, App
249	5	3.7	33	3	US-08-871-355A-312	Sequence 312, App	322	5	3.7	100	2	US-08-788-800-7	Sequence 7, App
250	5	3.7	33	4	US-09-201-945-312	Sequence 312, App	323	5	3.7	100	3	US-08-422-093-10	Sequence 10, App
251	5	3.7	34	4	US-08-833-752-13	Sequence 13, App	324	5	3.7	100	3	US-08-422-112-10	Sequence 10, App
252	5	3.7	35	2	US-08-455-625-1	Sequence 1, App	325	5	3.7	100	4	US-09-134-001C-4700	Sequence 4700, App
253	5	3.7	35	2	US-08-455-625-27	Sequence 27, App	326	5	3.7	102	2	US-08-480-473B-50	Sequence 50, App
254	5	3.7	35	4	US-08-455-685-1	Sequence 1, App	327	5	3.7	102	3	US-08-915-213-50	Sequence 50, App
255	5	3.7	35	4	US-08-455-685-27	Sequence 27, App	328	5	3.7	102	4	US-09-235-217-50	Sequence 50, App
256	5	3.7	35	4	US-08-905-223-331	Sequence 331, App	329	5	3.7	103	4	US-08-905-223-395	Sequence 395, App
257	5	3.7	35	4	US-08-060-988A-1	Sequence 1, App	330	5	3.7	105	4	US-09-187-789-26	Sequence 26, App
258	5	3.7	35	4	US-08-060-988A-27	Sequence 27, App	331	5	3.7	105	4	US-09-139-600-21	Sequence 21, App
259	5	3.7	35	4	PCT-US94-05142-1	Sequence 1, App	332	5	3.7	106	4	US-09-113-977C-49	Sequence 49, App
260	5	3.7	35	5	PCT-US94-05142-27	Sequence 27, App	333	5	3.7	107	1	US-08-216-852-87	Sequence 87, App
261	5	3.7	38	1	US-08-318-193-81	Sequence 81, App	334	5	3.7	107	1	US-08-899-575-87	Sequence 87, App
262	5	3.7	38	2	US-08-629-291A-21	Sequence 21, App	335	5	3.7	107	1	US-08-899-575-87	Sequence 87, App
263	5	3.7	38	2	US-08-658-335B-21	Sequence 21, App	336	5	3.7	107	5	PCT-US95-08724-87	Sequence 87, App
264	5	3.7	38	4	US-09-406-640-21	Sequence 21, App	337	5	3.7	108	1	US-08-485-359-2	Sequence 2, App
265	5	3.7	40	1	US-07-917-630-76	Sequence 76, App	338	5	3.7	108	1	US-08-569-594-2	Sequence 2, App
266	5	3.7	40	3	US-08-948-782-3	Sequence 3, App	339	5	3.7	108	1	US-08-468-661-3	Sequence 3, App
267	5	3.7	40	4	US-08-945-983-6	Sequence 6, App	340	5	3.7	108	1	US-08-466-772A-3	Sequence 3, App
268	5	3.7	40	4	US-09-482-612-3	Sequence 3, App	341	5	3.7	108	1	US-08-478-857-3	Sequence 3, App
269	5	3.7	40	4	US-09-677-554-3	Sequence 3, App	342	5	3.7	108	2	US-08-471-771-3	Sequence 3, App
270	5	3.7	43	1	US-07-956-700B-66	Sequence 66, App	343	5	3.7	108	3	US-09-130-783-3	Sequence 3, App
271	5	3.7	43	1	US-08-476-537-66	Sequence 66, App	344	5	3.7	108	5	PCT-US96-08815-2	Sequence 2, App
272	5	3.7	43	1	US-08-485-607-66	Sequence 66, App	345	5	3.7	109	1	US-08-485-359-4	Sequence 4, App
273	5	3.7	43	2	US-08-475-879-66	Sequence 66, App	346	5	3.7	109	1	US-08-569-594-4	Sequence 4, App
274	5	3.7	44	1	US-09-433-043B-66	Sequence 66, App	347	5	3.7	109	4	US-09-091-725-31	Sequence 31, App
275	5	3.7	44	2	US-08-262-037-131	Sequence 131, App	348	5	3.7	109	5	PCT-US96-08815-4	Sequence 4, App
276	5	3.7	44	1	US-08-262-037-132	Sequence 132, App	349	5	3.7	111	1	US-08-466-886-19	Sequence 19, App
277	5	3.7	44	1	US-08-894-483-5	Sequence 5, App	350	5	3.7	111	4	US-08-469-617-19	Sequence 19, App
278	5	3.7	45	4	US-08-900-230-43	Sequence 43, App	351	5	3.7	112	3	US-08-483-749A-9	Sequence 9, App
279	5	3.7	45	4	US-08-960-230-44	Sequence 44, App	352	5	3.7	112	3	US-08-928-361B-10	Sequence 10, App
280	5	3.7	45	4	US-08-469-260A-199	Sequence 199, App	353	5	3.7	112	3	US-08-928-361B-29	Sequence 29, App
281	5	3.7	46	4	US-08-900-230-40	Sequence 40, App	354	5	3.7	112	4	US-09-134-001C-3352	Sequence 3352, App
282	5	3.7	48	2	US-08-637-759B-480	Sequence 480, App	355	5	3.7	113	4	US-08-836-075A-92	Sequence 92, App
283	5	3.7	48	3	US-08-871-355A-480	Sequence 480, App	356	5	3.7	113	4	US-09-220-528-3	Sequence 3, App
284	5	3.7	48	4	US-09-201-945-480	Sequence 480, App	357	5	3.7	113	4	US-09-220-528-34	Sequence 34, App
285	5	3.7	52	2	US-08-480-473B-44	Sequence 44, App	358	5	3.7	113	4	US-09-406-532-14	Sequence 14, App
286	5	3.7	52	3	US-08-915-213-44	Sequence 44, App	359	5	3.7	116	4	US-08-961-536-4	Sequence 4, App
287	5	3.7	52	4	US-09-235-217-44	Sequence 44, App	360	5	3.7	116	4	US-09-220-528-4	Sequence 4, App
288	5	3.7	53	2	US-08-640-847C-40	Sequence 40, App	361	5	3.7	116	4	US-09-220-528-35	Sequence 35, App
289	5	3.7	53	3	US-08-872-094-2	Sequence 2, App	362	5	3.7	118	1	US-08-300-903A-13	Sequence 13, App
290	5	3.7	57	4	US-09-227-357-534	Sequence 534, App	363	5	3.7	118	2	US-08-476-866-24	Sequence 24, App
291	5	3.7	58	4	US-09-227-357-203	Sequence 203, App	364	5	3.7	120	2	US-08-652-558-4	Sequence 4, App
292	5	3.7	60	4	US-09-134-001C-4513	Sequence 4513, App	365	5	3.7	120	2	US-08-652-558-5	Sequence 5, App
293	5	3.7	62	3	US-08-894-483-8	Sequence 8, App	366	5	3.7	120	2	US-08-652-558-6	Sequence 6, App
294	5	3.7	67	3	US-09-268-070-3	Sequence 3, App	367	5	3.7	120	2	US-08-652-558-7	Sequence 7, App
295	5	3.7	69	4	US-09-710-099-14	Sequence 14, App	368	5	3.7	120	2	US-08-652-558-8	Sequence 8, App
296	5	3.7	73	4	US-09-134-001C-5108	Sequence 5108, App	369	5	3.7	120	2	US-08-652-558-9	Sequence 9, App
297	5	3.7	79	4	US-09-134-001C-3226	Sequence 3226, App	370	5	3.7	120	2	US-08-652-558-38	Sequence 38, App
298	5	3.7	80	1	US-08-011-398B-17	Sequence 17, App	371	5	3.7	120	4	US-09-254-189-2	Sequence 2, App
299	5	3.7	80	1	US-08-464-051-17	Sequence 17, App	372	5	3.7	120	4	US-09-254-189-3	Sequence 3, App
300	5	3.7	80	2	US-08-462-498-17	Sequence 17, App	373	5	3.7	120	4	US-09-254-189-4	Sequence 4, App
301	5	3.7	80	3	US-08-554-385-16	Sequence 16, App	374	5	3.7	120	4	US-09-254-189-5	Sequence 5, App
302	5	3.7	80	4	US-09-134-001C-5192	Sequence 5192, App	375	5	3.7	120	4	US-09-254-189-6	Sequence 6, App
303	5	3.7	84	1	US-08-279-590A-4	Sequence 4, App	376	5	3.7	120	4	US-09-336-536-45	Sequence 45, App
304	5	3.7	84	2	US-08-910-092-4	Sequence 4, App	377	5	3.7	120	4	US-09-149-476-426	Sequence 426, App
305	5	3.7	84	2	US-08-901-306-4	Sequence 4, App	378	5	3.7	124	4	US-09-199-637A-53	Sequence 53, App
306	5	3.7	84	4	US-09-134-001C-5031	Sequence 5031, App	379	5	3.7	124	4	US-08-387-805-12	Sequence 12, App
307	5	3.7	90	4	US-09-220-528-75	Sequence 75, App	380	5	3.7	126	4	US-09-228-986-100	Sequence 100, App
308	5	3.7	91	2	US-09-047-125-10	Sequence 10, App	381	5	3.7	126	4	US-09-342-647-16	Sequence 16, App
309	5	3.7	91	3	US-07-736-335E-10	Sequence 10, App	382	5	3.7	126	5	PCT-US96-03916-15	Sequence 15, App
310	5	3.7	91	4	US-09-342-647-8	Sequence 8, App	383	5	3.7	126	5	US-08-168-846-13	Sequence 13, App
311	5	3.7	92	1	US-08-347-492B-7	Sequence 7, App	384	5	3.7	132	4	US-08-125-642C-15	Sequence 15, App
312	5	3.7	92	2	US-08-798-143-7	Sequence 7, App	385	5	3.7	132	4	US-08-253-877C-10	Sequence 10, App
313	5	3.7	92	3	US-08-993-359-32	Sequence 32, App	386	5	3.7	133	1	US-08-452-164A-10	Sequence 10, App
314	5	3.7	93	4	US-08-981-392-44	Sequence 44, App	387	5	3.7	133	3	US-08-603-024-4	Sequence 4, App
315	5	3.7	94	1	US-08-591-498-4	Sequence 4, App	388	5	3.7	139	2	US-08-805-018-2	Sequence 2, App
316	5	3.7	96	4	US-09-230-528-19	Sequence 19, App	389	5	3.7	140	4	US-09-320-528-5	Sequence 5, App
317	5	3.7	96	4	US-09-220-528-33	Sequence 33, App	390	5	3.7	141	1	US-08-438-123-7	Sequence 7, App
318	5	3.7	96	4	US-09-057-889-41	Sequence 41, App	391	5	3.7	143	4	US-08-729-004-2	Sequence 2, App
319	5	3.7	100	1	US-08-422-101-10	Sequence 10, App	392	5	3.7	143	4	US-09-368-613-2	Sequence 2, App

393	5	3.7	144	4	US-09-220-528-36	Sequence 36, Appl	466	5	3.7	177	4	US-09-417-455-11	Sequence 11, Appl
394	5	3.7	145	3	US-09-030-613-9	Sequence 9, Appl1	467	5	3.7	177	4	US-09-348-942-11	Sequence 11, Appl
395	5	3.7	145	3	US-09-451-905-9	Sequence 9, Appl1	467	5	3.7	177	4	US-09-316-081-7	Sequence 7, Appl
396	5	3.7	149	2	US-08-460-694-5	Sequence 5, Appl1	468	5	3.7	177	4	US-09-578-458-7	Sequence 7, Appl
397	5	3.7	149	3	US-08-460-744-5	Sequence 5, Appl1	469	5	3.7	177	4	US-09-522-964-7	Sequence 7, Appl1
398	5	3.7	149	3	US-07-667-711B-5	Sequence 5, Appl1	470	5	3.7	177	4	US-09-626-11	Sequence 11, Appl
399	5	3.7	152	4	US-08-936-165A-64	Sequence 264, App	471	5	3.7	177	4	PCT-US96-03916-19	Sequence 19, Appl
400	5	3.7	152	4	US-08-100-744-3	Sequence 3, Appl1	472	5	3.7	177	4	US-08-680-326-4	Sequence 4, Appl1
401	5	3.7	153	1	US-08-854-811-3	Sequence 3, Appl1	473	5	3.7	178	2	US-09-134-001C-3756	Sequence 3756, Ap
402	5	3.7	153	2	US-08-284-784-3	Sequence 9, Appl1	474	5	3.7	179	1	US-07-668-648-8	Sequence 8, Appl1
403	5	3.7	154	1	US-08-463-262A-9	Sequence 9, Appl1	475	5	3.7	179	2	US-08-429-998-8	Sequence 8, Appl1
404	5	3.7	154	1	US-08-463-989-9	Sequence 67, Appl	476	5	3.7	179	2	US-08-431-333-8	Sequence 8, Appl1
405	5	3.7	154	1	US-08-450-945-67	Sequence 4, Appl1	477	5	3.7	179	2	US-09-125-619-18	Sequence 4237, Ap
406	5	3.7	154	4	US-08-964-652-4	Sequence 67, Appl	478	5	3.7	179	4	US-09-134-001C-4237	Sequence 18, Appl
407	5	3.7	154	4	US-08-976-161-67	Sequence 9, Appl1	479	5	3.7	179	4	US-09-366-862-2	Sequence 8, Appl1
408	5	3.7	154	4	US-09-003-574-9	Sequence 9, Appl1	480	5	3.7	179	5	PCT-US91-02321-8	Sequence 8, Appl1
409	5	3.7	154	4	US-09-003-570-9	Sequence 3132, Ap	481	5	3.7	181	4	US-08-505-187-3	Sequence 18, Appl
410	5	3.7	154	4	US-09-134-001C-3132	Sequence 298, App	482	5	3.7	181	4	US-09-220-528-40	Sequence 40, Appl
411	5	3.7	155	4	US-09-615-192A-298	Sequence 12, Appl	483	5	3.7	183	4	US-09-831-339-2	Sequence 50, Appl
412	5	3.7	158	4	US-08-612-973-12	Sequence 12, Appl	484	5	3.7	183	4	US-09-366-862-2	Sequence 2, Appl1
413	5	3.7	158	4	US-08-927-597-12	Sequence 21, Appl	485	5	3.7	183	4	US-09-382-276-1	Sequence 2, Appl1
414	5	3.7	158	4	US-09-125-619-21	Sequence 1, Appl1	486	5	3.7	183	4	US-09-382-276-1	Sequence 1, Appl1
415	5	3.7	159	1	US-08-193-182-1	Sequence 1, Appl1	487	5	3.7	183	4	PCT-US93-05240-14	Sequence 14, Appl
416	5	3.7	159	1	US-08-139-862-1	Sequence 2, Appl1	488	5	3.7	183	5	US-08-464-339A-2	Sequence 2, Appl1
417	5	3.7	159	1	US-08-139-862-2	Sequence 1, Appl1	489	5	3.7	183	6	US-08-468-847B-18	Sequence 18, Appl
418	5	3.7	159	2	US-08-829-110-1	Sequence 9, Appl1	490	5	3.7	184	1	US-08-531-525-40	Sequence 40, Appl
419	5	3.7	159	2	US-08-599-895-9	Sequence 9, Appl1	491	5	3.7	184	2	US-08-718-270A-40	Sequence 40, Appl
420	5	3.7	159	3	US-09-211-290-9	Sequence 9, Appl1	492	5	3.7	184	2	PCT-US94-14388-2	Sequence 2, Appl1
421	5	3.7	159	3	US-09-322-676-9	Sequence 10, Appl	493	5	3.7	185	4	US-08-612-973-24	Sequence 24, Appl
422	5	3.7	159	4	US-08-612-973-10	Sequence 9, Appl1	494	5	3.7	185	4	US-09-220-528-41	Sequence 41, Appl
423	5	3.7	159	4	US-08-927-597-10	Sequence 12, Appl	495	5	3.7	185	4	US-09-475-316A-13	Sequence 13, Appl
424	5	3.7	159	4	US-09-466-036A-9	Sequence 89, Appl	496	5	3.7	185	4	US-09-134-001C-5460	Sequence 5460, Ap
425	5	3.7	159	4	US-09-230-528-12	Sequence 89, Appl	497	5	3.7	185	4	US-09-134-001C-4095	Sequence 4095, Ap
426	5	3.7	159	6	5342615-3	Patent No. 5342615	498	5	3.7	186	4	US-09-382-276-2	Sequence 2, Appl1
427	5	3.7	159	6	5494663-6	Patent No. 5494663	499	5	3.7	186	4	US-09-125-619-23	Sequence 23, Appl1
428	5	3.7	161	4	US-08-858-207A-513	Sequence 513, App	500	5	3.7	186	4	US-09-206-935-16	Sequence 16, Appl
429	5	3.7	161	4	US-08-483-695-32	Sequence 32, Appl	501	5	3.7	187	4	US-07-681-701-17	Sequence 17, Appl
430	5	3.7	166	2	US-08-483-695-32	Sequence 32, Appl	502	5	3.7	187	4	US-08-086-428B-66	Sequence 66, Appl
431	5	3.7	166	2	US-07-965-285-32	Sequence 32, Appl	503	5	3.7	187	4	US-08-086-428B-66	Sequence 66, Appl
432	5	3.7	166	2	US-07-965-285-34	Sequence 34, Appl	504	5	3.7	187	4	US-08-086-428B-66	Sequence 66, Appl
433	5	3.7	166	2	US-08-487-231-32	Sequence 32, Appl	505	5	3.7	189	1	US-09-206-935-16	Sequence 16, Appl
434	5	3.7	166	2	US-08-487-231-32	Sequence 34, Appl	506	5	3.7	189	4	US-09-206-935-16	Sequence 16, Appl
435	5	3.7	166	2	US-08-487-231-32	Sequence 34, Appl	507	5	3.7	189	4	US-09-206-935-16	Sequence 16, Appl
436	5	3.7	166	2	US-08-487-231-32	Sequence 34, Appl	508	5	3.7	189	4	US-09-206-935-16	Sequence 16, Appl
437	5	3.7	166	4	US-09-201-912-34	Sequence 34, Appl	509	5	3.7	190	1	US-07-681-701-17	Sequence 17, Appl
438	5	3.7	166	4	US-09-339-913B-78	Sequence 34, Appl	510	5	3.7	192	1	US-08-086-428B-61	Sequence 61, Appl
439	5	3.7	166	4	US-09-339-913B-78	Sequence 34, Appl	511	5	3.7	192	1	US-08-086-428B-61	Sequence 61, Appl
440	5	3.7	166	4	US-09-339-913B-78	Sequence 34, Appl	512	5	3.7	192	1	US-08-086-428B-61	Sequence 61, Appl
441	5	3.7	166	4	US-09-339-913B-78	Sequence 34, Appl	513	5	3.7	192	1	US-08-086-428B-61	Sequence 61, Appl
442	5	3.7	166	4	US-09-339-913B-78	Sequence 34, Appl	514	5	3.7	192	1	US-08-086-428B-61	Sequence 61, Appl
443	5	3.7	166	4	US-09-339-913B-78	Sequence 34, Appl	515	5	3.7	192	1	US-08-086-428B-61	Sequence 61, Appl
444	5	3.7	166	4	US-09-339-913B-78	Sequence 34, Appl	516	5	3.7	192	1	US-08-086-428B-61	Sequence 61, Appl
445	5	3.7	167	4	US-09-224-110-2	Sequence 2, Appl1	517	5	3.7	192	1	US-08-086-428B-61	Sequence 61, Appl
446	5	3.7	167	4	US-09-224-110-2	Sequence 2, Appl1	518	5	3.7	192	1	US-08-086-428B-61	Sequence 61, Appl
447	5	3.7	167	5	PCr-US95-07289-2	Sequence 4508, Ap	519	5	3.7	192	1	US-08-086-428B-61	Sequence 61, Appl
448	5	3.7	168	1	US-08-362-453-17	Sequence 2, Appl1	520	5	3.7	192	1	US-08-086-428B-61	Sequence 61, Appl
449	5	3.7	168	1	US-08-441-139-10	Sequence 17, Appl	521	5	3.7	192	1	US-08-086-428B-61	Sequence 61, Appl
450	5	3.7	168	3	US-09-053-197A-20	Sequence 20, Appl	522	5	3.7	192	1	US-08-086-428B-61	Sequence 61, Appl
451	5	3.7	168	3	US-09-053-197A-20	Sequence 20, Appl	523	5	3.7	192	1	US-08-086-428B-61	Sequence 61, Appl
452	5	3.7	169	4	US-09-085-761A-20	Sequence 20, Appl	524	5	3.7	192	1	US-08-086-428B-61	Sequence 61, Appl
453	5	3.7	169	4	US-09-342-084-6	Sequence 20, Appl	525	5	3.7	192	1	US-08-086-428B-61	Sequence 61, Appl
454	5	3.7	169	4	US-09-134-001C-5390	Sequence 6, Appl1	526	5	3.7	192	2	US-08-468-570-66	Sequence 66, Appl
455	5	3.7	170	4	US-09-125-619-16	Sequence 16, Appl	527	5	3.7	192	2	US-08-468-570-66	Sequence 66, Appl
456	5	3.7	171	4	US-09-522-433B-24	Sequence 24, Appl	528	5	3.7	192	2	US-08-468-570-66	Sequence 66, Appl
457	5	3.7	172	4	US-09-173-151A-18	Sequence 18, Appl	529	5	3.7	192	2	US-08-468-570-66	Sequence 66, Appl
458	5	3.7	174	3	US-09-134-001C-2971	Sequence 2971, Ap	530	5	3.7	192	2	US-08-468-570-66	Sequence 66, Appl
459	5	3.7	174	3	US-08-480-173A-45	Sequence 45, Appl	531	5	3.7	192	2	US-08-468-570-66	Sequence 66, Appl
460	5	3.7	174	6	US-08-484-408A-45	Sequence 45, Appl	532	5	3.7	192	2	US-08-468-570-66	Sequence 66, Appl
461	5	3.7	176	1	5204096-2	Patent No. 5204096	533	5	3.7	192	2	US-08-468-570-66	Sequence 66, Appl
462	5	3.7	176	3	US-08-415-751-1	Sequence 1, Appl1	534	5	3.7	192	2	US-08-468-570-66	Sequence 66, Appl
463	5	3.7	176	3	US-09-000-630C-4	Sequence 4, Appl1	535	5	3.7	192	2	US-08-468-570-66	Sequence 66, Appl
464	5	3.7	176	3	US-08-862-730C-4	Sequence 4, Appl1	536	5	3.7	192	2	US-08-468-570-66	Sequence 66, Appl
465	5	3.7	177	3	US-08-469-260A-54	Sequence 54, Appl	537	5	3.7	192	2	US-08-468-570-66	Sequence 66, Appl
					US-09-000-630C-22	Sequence 22, Appl	538	5	3.7	192	2	US-08-468-570-66	Sequence 66, Appl
					US-08-862-730C-22	Sequence 22, Appl							

539	5	3.7	192	2	US-08-290-665A-71	Sequence 71, Appl	612	5	3.7	231	4	US-09-540-824-3	Sequence 3, Appl1
540	5	3.7	192	2	US-08-290-665A-72	Sequence 72, Appl	613	5	3.7	234	1	US-08-300-903A-11	Sequence 11, Appl1
541	5	3.7	192	2	US-08-290-665A-73	Sequence 73, Appl	614	5	3.7	235	1	US-07-932-454A-1	Sequence 1, Appl1
542	5	3.7	192	2	US-08-290-665A-75	Sequence 75, Appl	615	5	3.7	235	1	US-08-287-959-8	Sequence 8, Appl1
543	5	3.7	192	2	US-08-290-665A-76	Sequence 76, Appl	616	5	3.7	235	2	US-08-924-759-10	Sequence 10, Appl1
544	5	3.7	192	4	US-09-046-604-42	Sequence 42, Appl	617	5	3.7	235	3	US-09-248-335-10	Sequence 10, Appl
545	5	3.7	192	4	PCT-US95-10398-61	Sequence 61, Appl	618	5	3.7	236	1	US-08-112-208C-11	Sequence 11, Appl
546	5	3.7	192	5	PCT-US95-10398-63	Sequence 63, Appl	619	5	3.7	236	1	US-08-248-819A-11	Sequence 11, Appl
547	5	3.7	192	5	PCT-US95-10398-66	Sequence 66, Appl	620	5	3.7	236	1	US-08-607-269-22	Sequence 22, Appl
548	5	3.7	192	5	PCT-US95-10398-68	Sequence 68, Appl	621	5	3.7	236	1	US-08-442-063A-42	Sequence 42, Appl
549	5	3.7	192	5	PCT-US95-10398-70	Sequence 70, Appl	622	5	3.7	236	2	US-08-337-646A-11	Sequence 11, Appl
550	5	3.7	192	5	PCT-US95-10398-71	Sequence 71, Appl	623	5	3.7	236	2	US-08-856-531-11	Sequence 11, Appl
551	5	3.7	192	5	PCT-US95-10398-72	Sequence 72, Appl	624	5	3.7	236	2	US-08-856-034-11	Sequence 11, Appl
552	5	3.7	192	5	PCT-US95-10398-73	Sequence 73, Appl	625	5	3.7	236	4	US-09-127-048-9	Sequence 9, Appl1
553	5	3.7	192	5	PCT-US95-10398-75	Sequence 75, Appl	626	5	3.7	236	4	US-08-927-326-11	Sequence 11, Appl
554	5	3.7	192	5	PCT-US95-10398-76	Sequence 76, Appl	627	5	3.7	236	4	PCT-US95-04600-22	Sequence 22, Appl
555	5	3.7	194	1	US-08-117-083-21	Sequence 21, Appl	628	5	3.7	237	2	US-08-886-633-12	Sequence 2, Appl1
556	5	3.7	195	2	US-08-928-692-27	Sequence 27, Appl	629	5	3.7	237	3	US-08-470-535-14	Sequence 14, Appl
557	5	3.7	195	4	US-09-475-316A-106	Sequence 106, App	630	5	3.7	237	4	US-09-213-081-2	Sequence 2, Appl1
558	5	3.7	195	4	US-09-339-972-27	Sequence 27, Appl	631	5	3.7	237	4	US-09-212-979-2	Sequence 2, Appl1
559	5	3.7	195	4	US-09-288-143-144	Sequence 144, App	632	5	3.7	237	4	US-09-220-528-32	Sequence 32, Appl
560	5	3.7	195	4	US-09-125-619-17	Sequence 17, Appl	633	5	3.7	238	4	US-09-605-858-35	Sequence 35, Appl
561	5	3.7	197	4	US-09-112-248-2	Sequence 2, Appl1	634	5	3.7	238	4	US-08-858-207A-420	Sequence 420, App
562	5	3.7	197	4	US-08-975-215-4	Sequence 4, Appl1	635	5	3.7	239	4	US-08-933-750C-22	Sequence 22, Appl
563	5	3.7	197	4	US-09-776-271-4	Sequence 4, Appl1	636	5	3.7	239	4	US-09-234-613-22	Sequence 22, Appl
564	5	3.7	200	3	US-09-046-086-6	Sequence 6, Appl1	637	5	3.7	239	4	US-08-612-973-22	Sequence 22, Appl
565	5	3.7	200	4	US-08-612-973-26	Sequence 26, Appl	638	5	3.7	239	4	US-08-927-597-22	Sequence 22, Appl
566	5	3.7	200	4	US-08-927-597-26	Sequence 26, Appl	639	5	3.7	240	2	US-08-738-367-7	Sequence 7, Appl1
567	5	3.7	202	2	US-09-193-510-11	Sequence 11, Appl	640	5	3.7	240	2	US-09-410-464-12	Sequence 12, Appl
568	5	3.7	202	2	US-09-368-402-11	Sequence 11, Appl	641	5	3.7	243	1	US-08-021-608D-6	Sequence 6, Appl1
569	5	3.7	202	4	US-08-949-155-2	Sequence 2, Appl1	642	5	3.7	243	1	US-08-726-160-6	Sequence 6, Appl1
570	5	3.7	202	4	US-09-819-964-2	Sequence 2, Appl1	643	5	3.7	243	1	PCT-US94-01782-6	Sequence 6, Appl1
571	5	3.7	202	4	US-09-640-344-18	Sequence 18, Appl	644	5	3.7	244	1	US-07-869-933-32	Sequence 32, Appl
572	5	3.7	203	4	US-09-134-001C-4797	Sequence 4797, Ap	645	5	3.7	244	1	US-08-201-879A-3	Sequence 3, Appl1
573	5	3.7	207	4	US-09-068-960-43	Sequence 43, Appl	646	5	3.7	246	1	US-08-233-788A-41	Sequence 32, Appl
574	5	3.7	207	4	US-09-134-001C-3806	Sequence 3806, Ap	647	5	3.7	247	4	US-09-364-230-2	Sequence 41, Appl
575	5	3.7	208	4	US-08-680-726A-72	Sequence 72, Appl	648	5	3.7	247	4	US-09-399-913-20	Sequence 20, Appl
576	5	3.7	208	4	US-09-092-409-72	Sequence 72, Appl	649	5	3.7	247	4	US-09-399-913-22	Sequence 22, Appl
577	5	3.7	208	4	US-09-134-001C-4971	Sequence 4971, Ap	650	5	3.7	249	1	US-08-089-998B-2	Sequence 2, Appl1
578	5	3.7	209	1	US-07-892-827D-2	Sequence 2, Appl1	651	5	3.7	249	5	PCT-US94-07595-2	Sequence 2, Appl1
579	5	3.7	209	1	US-08-216-593-2	Sequence 2, Appl1	652	5	3.7	249	5	US-08-457-272-2	Sequence 2, Appl1
580	5	3.7	209	4	US-08-612-973-8	Sequence 8, Appl1	653	5	3.7	251	1	US-08-300-903A-7	Sequence 7, Appl1
581	5	3.7	209	4	US-08-927-597-8	Sequence 8, Appl1	654	5	3.7	252	4	US-09-399-913-20	Sequence 20, Appl
582	5	3.7	209	4	US-09-125-619-45	Sequence 45, Appl	655	5	3.7	252	4	US-09-399-913-22	Sequence 22, Appl
583	5	3.7	209	5	PCT-US93-12380-2	Sequence 2, Appl1	656	5	3.7	252	4	US-09-399-913-28	Sequence 28, Appl
584	5	3.7	210	4	US-08-612-973-28	Sequence 28, Appl	657	5	3.7	252	4	US-09-399-913-22	Sequence 22, Appl
585	5	3.7	210	4	US-08-927-597-28	Sequence 28, Appl	658	5	3.7	252	4	US-09-398-731-20	Sequence 20, Appl
586	5	3.7	210	4	US-08-612-973-4	Sequence 4, Appl1	659	5	3.7	252	4	US-09-298-731-22	Sequence 22, Appl
587	5	3.7	212	4	US-09-154-083-4	Sequence 4, Appl1	660	5	3.7	252	4	US-09-298-731-28	Sequence 28, Appl
588	5	3.7	212	4	US-08-927-597-4	Sequence 4, Appl1	661	5	3.7	254	2	US-08-407-481-20	Sequence 20, Appl
589	5	3.7	212	4	US-09-125-619-31	Sequence 31, Appl	662	5	3.7	254	4	US-09-202-548B-6	Sequence 6, Appl1
590	5	3.7	212	4	US-09-125-619-32	Sequence 32, Appl	663	5	3.7	254	5	PCT-US95-02689-20	Sequence 3, Appl1
591	5	3.7	212	4	US-09-125-619-43	Sequence 43, Appl	664	5	3.7	254	5	US-08-486-099-113	Sequence 20, Appl
592	5	3.7	212	4	US-09-125-619-44	Sequence 44, Appl	665	5	3.7	257	3	US-08-485-551A-113	Sequence 113, App
593	5	3.7	212	4	US-09-125-619-39	Sequence 39, Appl	666	5	3.7	257	3	US-08-471-913A-113	Sequence 113, App
594	5	3.7	213	4	US-09-125-619-35	Sequence 35, Appl	667	5	3.7	257	3	US-08-484-223B-113	Sequence 113, App
595	5	3.7	214	4	US-09-125-619-35	Sequence 35, Appl	668	5	3.7	257	3	US-08-485-264A-113	Sequence 113, App
596	5	3.7	215	4	US-09-087-232A-17	Sequence 17, Appl	669	5	3.7	257	3	US-08-475-668A-113	Sequence 113, App
597	5	3.7	215	4	US-08-833-752-6	Sequence 6, Appl1	670	5	3.7	257	3	US-08-475-668A-113	Sequence 113, App
598	5	3.7	216	1	US-08-041-774-5	Sequence 13, Appl1	671	5	3.7	257	3	US-08-471-913A-113	Sequence 113, App
599	5	3.7	216	2	US-08-562-985A-6	Sequence 6, Appl1	672	5	3.7	257	3	US-08-485-264A-113	Sequence 113, App
600	5	3.7	216	2	US-08-473-750-2	Sequence 2, Appl1	673	5	3.7	257	4	US-08-475-668A-113	Sequence 113, App
601	5	3.7	216	2	US-08-473-750-2	Sequence 2, Appl1	674	5	3.7	257	4	US-08-475-668A-113	Sequence 113, App
602	5	3.7	216	2	US-08-477-326-2	Sequence 2, Appl1	675	5	3.7	257	4	US-09-399-913-16	Sequence 16, Appl
603	5	3.7	216	4	US-08-530-340-5	Sequence 5, Appl1	676	5	3.7	257	4	US-09-298-731-16	Sequence 16, Appl
604	5	3.7	220	4	US-09-220-528-26	Sequence 26, Appl1	677	5	3.7	257	4	US-09-144-776B-2	Sequence 2, Appl1
605	5	3.7	223	3	US-08-820-970-2	Sequence 2, Appl1	678	5	3.7	258	4	US-09-134-001C-5593	Sequence 5593, Ap
606	5	3.7	223	4	US-08-811-682-8	Sequence 8, Appl1	679	5	3.7	262	2	US-08-481-956A-8	Sequence 8, Appl1
607	5	3.7	223	4	US-09-134-001C-4407	Sequence 4407, Ap	680	5	3.7	262	2	US-08-629-291A-8	Sequence 8, Appl1
608	5	3.7	224	4	US-08-768-373-6	Sequence 6, Appl1	681	5	3.7	262	2	US-08-658-335S-8	Sequence 8, Appl1
609	5	3.7	224	4	US-09-220-528-29	Sequence 29, Appl	682	5	3.7	262	4	US-09-134-001C-5011	Sequence 5011, Ap
610	5	3.7	224	4	US-09-134-001C-4010	Sequence 4010, Ap	683	5	3.7	262	4	US-09-134-001C-5532	Sequence 5532, Ap
611	5	3.7	225	1	US-08-300-903A-14	Sequence 14, Appl	684	5	3.7	262	4	US-09-406-640-8	Sequence 8, Appl1

685	5	3.7	263	1	US-07-927-071-2	Sequence 2, Appl1	758	5	3.7	286	4	US-08-727-616A-4	Sequence 4, Appl1
686	5	3.7	263	1	US-08-407-544-2	Sequence 2, Appl1	759	5	3.7	286	4	US-09-025-769B-265	Sequence 265, App
687	5	3.7	263	4	US-08-612-973-6	Sequence 6, Appl1	760	5	3.7	286	4	US-09-025-769B-362	Sequence 362, App
688	5	3.7	263	4	US-08-927-597-6	Sequence 6, Appl1	761	5	3.7	286	4	PCT-US91-07506-16	Sequence 16, Appl
689	5	3.7	263	4	US-09-134-001C-4512	Sequence 4512, Ap	762	5	3.7	286	4	US-08-437-607A-2	Sequence 2, Appl1
690	5	3.7	264	1	US-08-482-271-3	Sequence 3, Appl1	763	5	3.7	287	2	US-08-437-607A-2	Sequence 4005, Ap
691	5	3.7	264	1	US-08-482-271-4	Sequence 4, Appl1	764	5	3.7	288	4	US-09-134-001C-4005	Sequence 11, Appl
692	5	3.7	264	2	US-08-854-811-45	Sequence 45, Appl	765	5	3.7	289	1	US-08-110-300A-11	Sequence 11, Appl
693	5	3.7	264	2	US-08-854-811-45	Sequence 8, Appl1	766	5	3.7	289	2	US-08-886-642-11	Sequence 11, Appl
694	5	3.7	264	3	US-08-719-697-8	Sequence 2, Appl1	767	5	3.7	289	5	PCT-US93-08041-11	Sequence 11, Appl
695	5	3.7	264	3	US-09-080-120A-2	Sequence 4, Appl1	768	5	3.7	290	1	US-08-437-245-7	Sequence 7, Appl1
696	5	3.7	264	3	US-08-727-616A-8	Sequence 2, Appl1	769	5	3.7	290	4	US-09-071-035-96	Sequence 96, Appl
697	5	3.7	264	4	US-09-322-484-1	Sequence 1, Appl1	770	5	3.7	291	1	US-08-468-847B-19	Sequence 19, Appl
698	5	3.7	264	4	US-09-089-062-1	Sequence 2, Appl1	771	5	3.7	291	3	US-09-080-120A-7	Sequence 7, Appl1
699	5	3.7	264	5	PCT-US95-08925-2	Sequence 2, Appl1	772	5	3.7	291	6	PCT-US95-08925-7	Sequence 7, Appl1
700	5	3.7	265	2	US-08-719-697-2	Sequence 11, App	773	5	3.7	291	6	5212074-5	Patent No. 5212074
701	5	3.7	265	2	US-08-719-697-2	Sequence 2, Appl1	774	5	3.7	292	6	5258287-24	Patent No. 5258287
702	5	3.7	265	2	US-08-719-697-6	Sequence 6, Appl1	775	5	3.7	293	2	US-08-919-145-2	Sequence 2, Appl1
703	5	3.7	265	4	US-08-727-616A-2	Sequence 2, Appl1	776	5	3.7	293	4	US-09-344-889-2	Sequence 2, Appl1
704	5	3.7	265	4	US-08-727-616A-6	Sequence 6, Appl1	777	5	3.7	295	4	US-08-952-089A-1	Sequence 1, Appl1
705	5	3.7	266	2	US-08-484-905-115	Sequence 115, App	778	5	3.7	295	4	US-09-199-637A-341	Sequence 341, App
706	5	3.7	266	3	US-08-481-985B-115	Sequence 115, App	779	5	3.7	299	4	US-09-025-765B-285	Sequence 285, App
707	5	3.7	266	3	US-09-147-550-115	Sequence 115, App	780	5	3.7	299	4	US-09-025-769B-298	Sequence 298, App
708	5	3.7	266	4	US-08-370-476-115	Sequence 115, App	781	5	3.7	299	4	US-09-025-769B-300	Sequence 300, App
709	5	3.7	266	4	US-09-557-917-11	Sequence 11, Appl	782	5	3.7	300	1	US-07-640-029-5	Sequence 5, Appl1
710	5	3.7	266	6	516941-1	Patent No. 516941	783	5	3.7	300	1	US-08-439-992A-5	Sequence 5, Appl1
711	5	3.7	267	1	US-08-300-903A-15	Sequence 15, Appl	784	5	3.7	301	1	US-08-674-168-31	Sequence 31, Appl
712	5	3.7	269	4	US-08-965-056-30	Sequence 30, Appl	785	5	3.7	301	3	US-08-985-908-15	Sequence 15, Appl
713	5	3.7	270	4	US-09-399-913-14	Sequence 14, Appl	786	5	3.7	301	4	US-08-856-935-5	Sequence 20, Appl
714	5	3.7	270	4	US-09-339-913-18	Sequence 18, Appl	787	5	3.7	302	1	US-07-640-029-6	Sequence 5, Appl1
715	5	3.7	270	4	US-09-298-731-14	Sequence 14, Appl	788	5	3.7	302	1	US-07-640-029-6	Sequence 6, Appl1
716	5	3.7	270	4	US-09-298-731-18	Sequence 18, Appl	789	5	3.7	302	1	US-07-921-807B-7	Sequence 7, Appl1
717	5	3.7	271	2	US-08-400-115-4	Sequence 4, Appl1	790	5	3.7	302	1	US-08-441-944A-8	Sequence 8, Appl1
718	5	3.7	271	2	US-08-599-895-7	Sequence 7, Appl1	791	5	3.7	302	1	US-08-441-944A-8	Sequence 8, Appl1
719	5	3.7	271	3	US-09-211-290-7	Sequence 7, Appl1	792	5	3.7	302	4	US-09-077-675A-2	Sequence 7, Appl1
720	5	3.7	271	3	US-09-030-613-7	Sequence 7, Appl1	793	5	3.7	302	4	US-09-077-675A-7	Sequence 7, Appl1
721	5	3.7	271	3	US-09-322-676-7	Sequence 7, Appl1	794	5	3.7	302	4	US-08-439-992A-6	Sequence 6, Appl1
722	5	3.7	271	4	US-09-085-305-7	Sequence 7, Appl1	795	5	3.7	302	4	US-09-282-305-14	Sequence 14, Appl
723	5	3.7	271	4	US-09-077-675A-12	Sequence 12, Appl	796	5	3.7	303	4	US-09-632-947B-5	Sequence 5, Appl1
724	5	3.7	271	4	US-09-466-036A-7	Sequence 7, Appl1	797	5	3.7	305	4	US-09-282-305-12	Sequence 12, Appl
725	5	3.7	271	4	US-09-451-905-7	Sequence 7, Appl1	798	5	3.7	306	1	US-08-317-522A-7	Sequence 7, Appl1
726	5	3.7	272	6	5494663-3	Patent No. 5494663	799	5	3.7	306	1	US-08-439-818A-7	Sequence 7, Appl1
727	5	3.7	272	4	US-09-134-001C-4238	Sequence 4238, Ap	800	5	3.7	306	2	US-08-751-965-7	Sequence 7, Appl1
728	5	3.7	273	4	US-09-147-915-4	Sequence 4, Appl1	801	5	3.7	306	2	US-08-738-975-7	Sequence 7, Appl1
729	5	3.7	278	1	US-08-188-582-9	Sequence 9, Appl1	802	5	3.7	306	2	US-08-728-626-7	Sequence 7, Appl1
730	5	3.7	278	3	US-08-646-715-9	Sequence 9, Appl1	803	5	3.7	306	3	US-08-808-599A-7	Sequence 7, Appl1
731	5	3.7	278	3	US-08-663-082-4	Sequence 4, Appl1	804	5	3.7	306	4	US-09-105-390-36	Sequence 7, Appl1
732	5	3.7	278	4	US-09-036-987A-22	Sequence 22, Appl	805	5	3.7	306	4	US-09-342-647-26	Sequence 26, Appl
733	5	3.7	278	4	US-09-370-700-22	Sequence 9, Appl1	806	5	3.7	307	4	US-08-442-063A-48	Sequence 4678, Ap
734	5	3.7	280	1	US-08-300-903A-9	Sequence 6, Appl1	807	5	3.7	307	4	US-08-975-215-2	Sequence 2, Appl1
735	5	3.7	280	1	US-08-595-559-3	Sequence 52, Appl	808	5	3.7	307	4	US-07-776-271-2	Sequence 2, Appl1
736	5	3.7	281	2	US-08-284-465-6	Sequence 6, Appl1	809	5	3.7	308	1	US-07-828-700-9	Sequence 8, Appl1
737	5	3.7	282	1	US-08-118-270-52	Sequence 52, Appl	810	5	3.7	309	2	US-08-463-092B-8	Sequence 2, Appl1
738	5	3.7	282	4	US-08-442-063A-45	Sequence 9, Appl1	811	5	3.7	309	2	US-08-460-907B-8	Sequence 8, Appl1
739	5	3.7	282	4	US-09-455-960-9	Sequence 52, Appl	812	5	3.7	310	1	US-08-433-783-42	Sequence 42, Appl
740	5	3.7	283	4	PCT-US93-08528-52	Sequence 1, Appl1	813	5	3.7	310	1	US-08-337-358-4	Sequence 42, Appl
741	5	3.7	283	4	US-09-367-012-1	Sequence 33, Appl	814	5	3.7	310	5	PCT-US95-07537A-42	Sequence 42, Appl
742	5	3.7	283	4	US-09-180-109A-30	Sequence 30, Appl	815	5	3.7	310	5	US-09-073-358-4	Sequence 42, Appl
743	5	3.7	283	4	US-09-180-109A-33	Sequence 33, Appl	816	5	3.7	310	5	US-09-124-758-2	Sequence 29, Appl
744	5	3.7	283	4	US-09-777-157A-1	Sequence 5, Appl1	817	5	3.7	313	4	US-09-342-647-29	Sequence 12, Appl
745	5	3.7	284	2	US-08-437-607A-5	Sequence 4, Appl1	818	5	3.7	313	4	US-09-202-548B-12	Sequence 94, Appl
746	5	3.7	285	4	US-09-412-102-4	Sequence 4, Appl1	819	5	3.7	315	4	US-08-827-231A-3	Sequence 2, Appl1
747	5	3.7	285	4	US-09-217-787-4	Sequence 2, Appl1	820	5	3.7	315	4	US-08-347-395A-2	Sequence 2, Appl1
748	5	3.7	285	4	US-09-134-001C-4455	Sequence 4455, Ap	821	5	3.7	316	2	US-07-992-827D-1	Sequence 17, Appl
749	5	3.7	286	1	US-07-721-775A-2	Sequence 2, Appl1	822	5	3.7	320	4	US-09-256-000-17	Sequence 10, Appl
750	5	3.7	286	1	US-08-339-658-2	Sequence 9, Appl1	823	5	3.7	323	1	US-08-110-300A-10	Sequence 10, Appl
751	5	3.7	286	1	US-08-246-403A-9	Sequence 12, Appl	824	5	3.7	323	1	US-08-216-593-1	Sequence 1, Appl1
752	5	3.7	286	1	US-08-346-403A-12	Sequence 16, Appl	825	5	3.7	323	1	US-08-540-804-18	Sequence 18, Appl
753	5	3.7	286	2	US-08-719-697-4	Sequence 7, Appl1	826	5	3.7	323	2	US-08-218-265-18	Sequence 18, Appl
754	5	3.7	286	4	US-09-263-933-7	Sequence 14, Appl	827	5	3.7	323	2	US-08-886-642-10	Sequence 10, Appl
755	5	3.7	286	4	US-09-263-933-14	Sequence 21, Appl	828	5	3.7	323	2	US-08-521-872-18	Sequence 18, Appl
756	5	3.7	286	4	US-09-263-933-21	Sequence 21, Appl	829	5	3.7	323	3		
757	5	3.7	286	4			830	5	3.7	323	3		

831	5	3.7	323	4	US-08-643-212-18	Sequence 18, Appl	904	5	3.7	349	2	US-08-465-971B-2	Sequence 2, Appl1
832	5	3.7	323	4	US-08-643-212-20	Sequence 20, Appl	905	5	3.7	351	2	US-08-933-750C-19	Sequence 19, Appl
833	5	3.7	323	4	US-08-643-212-22	Sequence 22, Appl	906	5	3.7	351	4	US-09-234-613-19	Sequence 19, Appl
834	5	3.7	323	4	US-08-643-212-24	Sequence 24, Appl	907	5	3.7	352	4	US-08-698-805-2	Sequence 3, Appl1
835	5	3.7	323	4	US-08-643-212-26	Sequence 26, Appl	908	5	3.7	353	4	US-09-077-675A-3	Sequence 3, Appl1
836	5	3.7	323	4	US-08-643-212-28	Sequence 28, Appl	909	5	3.7	354	2	US-07-868-353A-12	Sequence 12, Appl
837	5	3.7	323	4	US-08-590-399-18	Sequence 18, Appl	910	5	3.7	354	2	US-08-407-804-21	Sequence 21, Appl
838	5	3.7	323	4	PCT-US93-08041-10	Sequence 10, Appl	911	5	3.7	354	3	US-09-124-807-21	Sequence 21, Appl
839	5	3.7	323	5	PCT-US93-12380-1	Sequence 10, Appl	912	5	3.7	354	4	US-09-134-001C-3278	Sequence 3278, Ap
840	5	3.7	325	6	5320941-2	Patent No. 5320941	913	5	3.7	355	4	US-09-045-583-53	Sequence 53, Appl
841	5	3.7	326	4	US-09-632-947B-1	Sequence 1, Appl1	914	5	3.7	355	4	US-09-534-185-53	Sequence 53, Appl
842	5	3.7	327	1	US-08-118-270-55	Sequence 55, Appl	915	5	3.7	356	4	US-09-125-619-2	Sequence 2, Appl1
843	5	3.7	327	5	PCT-US93-08528-55	Sequence 55, Appl	916	5	3.7	356	4	US-09-125-619-2	Sequence 13, Appl1
844	5	3.7	329	2	US-08-913-477-2	Sequence 4, Appl1	917	5	3.7	356	6	5223606-7	Patent No. 5223606
845	5	3.7	329	2	US-08-913-477-4	Sequence 4, Appl1	918	5	3.7	357	1	US-08-078-683A-8	Sequence 8, Appl1
846	5	3.7	329	4	US-08-887-534A-24	Sequence 24, Appl	919	5	3.7	357	1	US-08-638-911A-37	Sequence 37, Appl
847	5	3.7	330	4	US-09-359-161-6	Sequence 24, Appl	920	5	3.7	358	1	US-08-453-117-4	Sequence 4, Appl1
848	5	3.7	331	2	US-08-385-191A-2	Sequence 2, Appl1	921	5	3.7	358	2	US-08-948-222-4	Sequence 4, Appl1
849	5	3.7	331	2	US-08-878-989-21	Sequence 21, Appl1	922	5	3.7	358	2	US-08-973-145-4	Sequence 3, Appl1
850	5	3.7	331	3	US-09-101-146-64	Sequence 64, Appl	923	5	3.7	358	2	US-08-165-971B-3	Sequence 3, Appl1
851	5	3.7	331	4	US-08-849-751-4	Sequence 4, Appl1	924	5	3.7	358	2	US-08-973-145-4	Sequence 3, Appl1
852	5	3.7	331	4	US-09-272-796-21	Sequence 21, Appl	925	5	3.7	359	1	PCT-US96-08081-4	Sequence 4, Appl1
853	5	3.7	331	4	US-09-478-816-4	Sequence 4, Appl1	926	5	3.7	359	1	US-08-303-238-4	Sequence 4, Appl1
854	5	3.7	331	4	US-08-472-402A-27	Sequence 2, Appl1	927	5	3.7	361	4	US-08-458-834-4	Sequence 4, Appl1
855	5	3.7	333	4	US-08-442-063A-27	Sequence 27, Appl	928	5	3.7	361	4	US-09-077-675A-8	Sequence 8, Appl1
856	5	3.7	333	1	US-08-117-083-64	Sequence 64, Appl	929	5	3.7	362	4	US-08-415-751-5	Sequence 5, Appl1
857	5	3.7	334	1	US-08-287-442-9	Sequence 9, Appl1	930	5	3.7	362	4	US-09-134-001C-4670	Sequence 4670, Ap
858	5	3.7	334	1	US-08-459-701-9	Sequence 9, Appl1	931	5	3.7	364	4	US-09-077-675A-16	Sequence 16, Appl
859	5	3.7	334	1	US-08-460-298-9	Sequence 9, Appl1	932	5	3.7	365	4	US-09-336-536-40	Sequence 40, Appl
860	5	3.7	334	1	US-08-459-174-9	Sequence 9, Appl1	933	5	3.7	366	4	US-09-077-675A-13	Sequence 13, Appl
861	5	3.7	334	1	US-09-105-390-52	Sequence 52, Appl	934	5	3.7	370	2	US-08-360-606B-32	Sequence 32, Appl
862	5	3.7	336	1	US-08-806-581A-2	Sequence 2, Appl1	935	5	3.7	371	4	US-09-134-001C-3829	Sequence 3829, Ap
863	5	3.7	338	4	US-08-961-536-2	Sequence 2, Appl1	936	5	3.7	372	1	US-08-186-218-33	Sequence 33, Appl
864	5	3.7	338	4	US-09-134-001C-5128	Sequence 5128, Ap	937	5	3.7	372	1	US-08-681-953-33	Sequence 33, Appl
865	5	3.7	339	4	US-09-232-191-2	Sequence 2, Appl1	938	5	3.7	373	4	US-09-039-198A-14	Sequence 14, Appl
866	5	3.7	339	4	US-09-232-200-2	Sequence 2, Appl1	939	5	3.7	373	4	US-08-877-599-14	Sequence 14, Appl
867	5	3.7	339	4	US-09-232-197-2	Sequence 2, Appl1	940	5	3.7	373	4	US-08-877-599-15	Sequence 15, Appl
868	5	3.7	339	4	US-09-232-201-2	Sequence 2, Appl1	941	5	3.7	373	4	US-08-267-574-14	Sequence 14, Appl
869	5	3.7	340	1	US-07-828-700-8	Sequence 8, Appl1	942	5	3.7	373	4	US-09-267-574-15	Sequence 15, Appl
870	5	3.7	340	1	US-08-462-195-2	Sequence 2, Appl1	943	5	3.7	375	1	US-08-121-714-5	Sequence 5, Appl1
871	5	3.7	340	2	US-08-636-883-2	Sequence 2, Appl1	944	5	3.7	375	1	US-08-477-108A-5	Sequence 5, Appl1
872	5	3.7	340	2	US-08-757-653-176	Sequence 176, App	945	5	3.7	375	2	US-08-477-112-5	Sequence 5, Appl1
873	5	3.7	340	2	US-08-446-873-16	Sequence 16, Appl	946	5	3.7	375	2	PCT-US93-08332-5	Sequence 5, Appl1
874	5	3.7	340	2	US-08-823-516-79	Sequence 79, Appl	947	5	3.7	376	1	US-08-279-590A-2	Sequence 2, Appl1
875	5	3.7	340	2	US-08-823-516-136	Sequence 136, App	948	5	3.7	376	2	US-08-910-092-2	Sequence 2, Appl1
876	5	3.7	340	3	US-09-127-829-2	Sequence 2, Appl1	949	5	3.7	377	4	US-09-134-001C-3200	Sequence 3200, Ap
877	5	3.7	340	3	US-08-759-038-115	Sequence 115, App	950	5	3.7	382	1	US-08-470-299-7	Sequence 7, Appl1
878	5	3.7	340	3	US-08-758-314-115	Sequence 115, App	951	5	3.7	382	1	US-08-470-299-10	Sequence 10, Appl
879	5	3.7	341	1	US-08-396-957A-5	Sequence 5, Appl	952	5	3.7	382	4	US-09-384-305-2	Sequence 2, Appl1
880	5	3.7	342	1	US-08-272-919-2	Sequence 2, Appl1	953	5	3.7	383	4	US-09-459-749D-17	Sequence 17, Appl
881	5	3.7	342	1	US-08-619-916-2	Sequence 2, Appl1	954	5	3.7	384	4	US-09-032-215-22	Sequence 22, Appl
882	5	3.7	342	2	US-08-845-295A-3	Sequence 3, Appl1	955	5	3.7	384	4	US-09-071-035-276	Sequence 276, App
883	5	3.7	342	2	US-09-140-833-3	Sequence 3, Appl1	956	5	3.7	385	2	US-08-694-915-2	Sequence 2, Appl1
884	5	3.7	342	3	US-08-852-824-2	Sequence 2, Appl1	957	5	3.7	387	1	US-08-258-261B-19	Sequence 19, Appl
885	5	3.7	342	4	US-09-146-661-3	Sequence 3, Appl1	958	5	3.7	387	1	US-08-456-837-19	Sequence 19, Appl
886	5	3.7	342	4	US-09-150-515-3	Sequence 3, Appl1	959	5	3.7	387	1	US-08-457-342-19	Sequence 19, Appl
887	5	3.7	342	4	US-09-632-947B-2	Sequence 2, Appl1	960	5	3.7	387	1	US-08-457-646A-19	Sequence 19, Appl
888	5	3.7	342	4	US-09-632-947B-10	Sequence 10, Appl	961	5	3.7	387	1	US-08-458-076A-19	Sequence 19, Appl
889	5	3.7	342	5	PCT-US95-08542-2	Sequence 2, Appl1	962	5	3.7	387	1	US-08-457-335A-19	Sequence 19, Appl
890	5	3.7	343	3	US-09-109-204-32	Sequence 3, Appl	963	5	3.7	387	2	US-08-729-214-19	Sequence 19, Appl
891	5	3.7	344	1	US-08-657-192-3	Sequence 3, Appl1	964	5	3.7	387	2	US-08-486-839-6	Sequence 6, Appl1
892	5	3.7	344	3	US-08-523-373-5	Sequence 5, Appl1	965	5	3.7	387	3	US-09-151-011-6	Sequence 6, Appl1
893	5	3.7	344	3	US-09-393-554-4	Sequence 4, Appl1	966	5	3.7	387	3	US-09-028-924-19	Sequence 19, Appl
894	5	3.7	344	4	US-09-134-001C-3524	Sequence 3524, Ap	967	5	3.7	387	4	US-09-343-623-6	Sequence 6, Appl1
895	5	3.7	345	4	US-09-364-230-8	Sequence 8, Appl	968	5	3.7	388	4	US-09-219-983A-29	Sequence 29, Appl1
896	5	3.7	347	4	US-09-188-930-326	Sequence 326, App	969	5	3.7	388	6	5240849-3	Patent No. 5240849
897	5	3.7	348	4	US-09-671-950-2	Sequence 4, Appl1	970	5	3.7	390	4	US-09-308-003-12	Sequence 12, Appl
898	5	3.7	348	4	US-09-671-950-4	Sequence 4, Appl1	971	5	3.7	390	4	US-09-134-001C-3858	Sequence 3858, Ap
899	5	3.7	348	4	US-09-671-950-6	Sequence 6, Appl1	972	5	3.7	391	3	US-08-968-563-11	Sequence 11, Appl
900	5	3.7	348	4	US-09-671-950-8	Sequence 8, Appl1	973	5	3.7	391	4	US-08-969-683A-11	Sequence 11, Appl
901	5	3.7	348	4	US-09-671-950-10	Sequence 10, Appl	974	5	3.7	391	4	US-09-297-928-7	Sequence 7, Appl1
902	5	3.7	348	4	US-09-671-950-12	Sequence 12, Appl	975	5	3.7	392	1	US-08-451-777A-33	Sequence 33, Appl
903	5	3.7	348	4	US-09-671-950-14	Sequence 14, Appl	976	5	3.7	392	2	US-08-451-778A-33	Sequence 33, Appl

977 5 3.7 392 2 US-08-998-208-33 Sequence 33, Appl
978 5 3.7 392 3 US-08-523-373-6 Sequence 6, Appl
979 5 3.7 393 4 US-09-260-843-2 Sequence 2, Appl
980 5 3.7 393 4 US-09-377-557-14 Sequence 14, Appl
981 5 3.7 393 4 US-09-194-905-13 Sequence 13, Appl
982 5 3.7 393 4 US-09-923-654-2 Sequence 2, Appl
983 5 3.7 394 2 US-08-646-5908-40 Sequence 40, Appl
984 5 3.7 394 4 US-09-412-184-40 Sequence 40, Appl
985 5 3.7 394 4 US-09-586-719-6 Sequence 6, Appl
986 5 3.7 394 4 US-09-134-001C-4954 Sequence 12, Appl
987 5 3.7 397 4 US-09-336-536-39 Sequence 39, Appl
988 5 3.7 397 4 US-09-036-987A-12 Sequence 12, Appl
989 5 3.7 397 4 US-09-370-700-12 Sequence 12, Appl
990 5 3.7 401 2 US-08-591-079-4 Sequence 4, Appl
991 5 3.7 401 2 US-08-591-079-6 Sequence 6, Appl
992 5 3.7 401 3 US-08-974-022-2 Sequence 2, Appl
993 5 3.7 401 4 US-08-795-445A-2 Sequence 2, Appl
994 5 3.7 401 4 US-08-795-445A-2 Sequence 2, Appl
995 5 3.7 401 4 US-08-795-447A-2 Sequence 2, Appl
996 5 3.7 401 4 US-08-974-186-2 Sequence 2, Appl
997 5 3.7 401 4 US-08-795-446B-2 Sequence 2, Appl
998 5 3.7 402 2 US-08-706-945D-124 Sequence 124, Appl
999 5 3.7 402 2 US-08-477-234A-2 Sequence 2, Appl
1000 5 3.7 402 2 US-08-472-576B-2 Sequence 2, Appl
PCT-US93-05705-14 Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-082-279B-1138
; Sequence 1138, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Metutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE REFERENCE: PHARMACOKINETIC PROPERTIES
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1138
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-1138

Query Match
Best Local Similarity 5.2%; Score 7; DB 4; Length 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 ANLTSL 39
Db 6 ANLTSL 12

RESULT 2
US-09-315-304B-1138
; Sequence 1138, Application US/09315304B
; Patent No. 6348568
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Metutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.

; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-052
; CURRENT APPLICATION NUMBER: US/09/315,304B
; CURRENT FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1138
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-315-304B-1138

Query Match
Best Local Similarity 5.2%; Score 7; DB 4; Length 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 3
US-08-707-873-1
; Sequence 1, Application US/08707873
; Patent No. 5747318
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL
; APPLICANT: DENG, TILIANG
; TITLE OF INVENTION: NOVEL MITOGEN ACTIVATED PROTEIN KINASE,
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESS: SPENSLEY HORN JUBAS & LOBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/707,873
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/315,067
; FILING DATE: 29-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: HAILE PH.D., LISA A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: PD3743
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..11
US-08-707-873-1

Query Match 4.5%; Score 6; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 ASTPES 26
|11111|
DB 5 ASTPES 10

RESULT 4
US-08-707-874-1

; Sequence 1, Application US/08707874
; Patent No. 5837451
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL
; APPLICANT: DENG, TILIANG
; TITLE OF INVENTION: NOVEL MITOGEN ACTIVATED PROTEIN KINASE,
; TITLE OF INVENTION: FRK
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/707,874
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/315,067
; FILING DATE: 29-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: HAILE PH.D., LISA A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: PD3743
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..11
US-08-707-874-1

Query Match 4.5%; Score 6; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 ASTPES 26
|11111|
DB 5 ASTPES 10

RESULT 5
US-08-315-067-1
; Sequence 1, Application US/08315067
; Patent No. 592557
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL

; APPLICANT: DENG, TILIANG
; TITLE OF INVENTION: NOVEL MITOGEN ACTIVATED PROTEIN KINASE,
; TITLE OF INVENTION: FRK
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90067

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/315,067
; FILING DATE: 29-SEP-1994
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:
; NAME: HAILE PH.D., LISA A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: PD3743
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..11
US-08-315-067-1

Query Match 4.5%; Score 6; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 ASTPES 26
|11111|
DB 5 ASTPES 10

RESULT 6
US-09-193-797-1

; Sequence 1, Application US/09193797
; Patent No. 6034560
; GENERAL INFORMATION:
; APPLICANT: KARIN, MICHAEL
; APPLICANT: DENG, TILIANG
; TITLE OF INVENTION: NOVEL MITOGEN ACTIVATED PROTEIN KINASE,
; TITLE OF INVENTION: FRK
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/193,797
; FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/315,067
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HAILE PH.D., LISA A.
REGISTRATION NUMBER: P-38,347
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/455-5100
TELEFAX: 619/455-5100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..11
US-09-193-797-1

Query Match 4.5%; Score 6; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 ASTPES 26
Db 5 ASTPES 10

RESULT 7
PCT-US95-01770-1
Sequence 1, Application PC/TUS9501770
GENERAL INFORMATION:
APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
TITLE OF INVENTION: NOVEL MITOGEN ACTIVATED PROTEIN KINASE,
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROBBINS, BERLINER & CARSON
STREET: 201 N. FIGUEROA STREET, 5TH FLOOR
CITY: LOS ANGELES
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 90012-2628
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01770
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: BERLINER, ROBERT
REGISTRATION NUMBER: 20,121
REFERENCE/DOCKET NUMBER: 5555-297
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213-977-1001
TELEFAX: 213-977-1003
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..11

PCT-US95-01770-1

Query Match 4.5%; Score 6; DB 5; Length 11;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 ASTPES 26
Db 5 ASTPES 10

RESULT 8

US-08-487-795A-24
Sequence 24, Application US/08487795A
Patent No. 6436390
GENERAL INFORMATION:
APPLICANT: Tekamp-Olson, Patricia
APPLICANT: Mullenbach, Guy
APPLICANT: Wernette-Hammond, Mary Ellen
TITLE OF INVENTION: IL8 INHIBITORS
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street R-440
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,795A
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/121,105
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Chung, Ling-Fong
REGISTRATION NUMBER: 36,482
REFERENCE/DOCKET NUMBER: 0949, 002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 923-2704
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-487-795A-24

US-08-487-795A-24

Query Match 4.5%; Score 6; DB 4; Length 39;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 PCMLET 123
Db 29 PCMLET 34

RESULT 9
US-08-121-105B-24

Sequence 24, Application US/08121105B
Patent No. 6448379
GENERAL INFORMATION:
APPLICANT: Tekamp-Olson, Patricia
APPLICANT: Mullenbach, Guy
APPLICANT: Wernette-Hammond, Mary Ellen

TITLE OF INVENTION: IL8 INHIBITORS
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street R-440
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/121,105B
FILING DATE: 14-SEP-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Chung, Ling-Fong
REGISTRATION NUMBER: 36,482
REFERENCE/DOCKET NUMBER: 0949.001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 923-2704
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-121-105B-24

Query Match 4.5%; Score 6; DB 4; Length 39;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 PCMLET 123
|||||
Db 29 PCMLET 34

RESULT 10
PCT-US94-10356-27
Sequence 27, Application PC/TUS9410356
GENERAL INFORMATION:
APPLICANT: Tekamp-Olson, Patricia Mullenbach,
APPLICANT: Guy Wernette-Hammond, Mary Ellen
TITLE OF INVENTION: IL8 INHIBITORS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street R-440
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10356
FILING DATE: 13-SEP-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/121,105
FILING DATE: 14-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Chung, Ling-Fong
REGISTRATION NUMBER: 36,482

REFERENCE/DOCKET NUMBER: 0949.100
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2704
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US94-10356-27

Query Match 4.5%; Score 6; DB 5; Length 39;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 PCMLET 123
|||||
Db 29 PCMLET 34

RESULT 11
US-08-469-537A-38
Sequence 38, Application US/08469537A
Patent No. 5843749
GENERAL INFORMATION:
APPLICANT: Maisongier, et al.
TITLE OF INVENTION: EHK AND ROR TYROSINE
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: NY
COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: USSN 08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: USSN 07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kempner, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-469-537A-38

Query Match 4.5%; Score 6; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 68;

Matches 6: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Oy 91 LTELTT 96
|||||

Db 71 LTELTT 76

RESULT 12
5169835-26
Patent No. 5169835
APPLICANT: MAI-YEE, CHAN
TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS
NUMBER OF SEQUENCES: 48
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/390,409
FILING DATE: 07-AUG-1989
SEQ ID NO: 26:
LENGTH: 93
5169835-26

Query Match
Best Local Similarity 4.5%; Score 6; DB 6; Length 93;
100.0%; Pred. No. 82;
Matches 6: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Oy 87 EAVSLT 92
|||||

Db 20 EAVSLT 25

RESULT 13
US-08-750-856A-16
Sequence 16, Application US/08750856A
Patent No. 5858672
GENERAL INFORMATION:
APPLICANT: SONIGO, PIERRE
APPLICANT: PANCINO, GIANFRANCO
APPLICANT: PETERHANS, ERNST
APPLICANT: BERTONI, GIUSEPPE
TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS AND CORRESPONDING
TITLE OF INVENTION: PEPTIDE FRAGMENTS FROM THE CAPRINE ARTHRITIS-ENCEPHALITIS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/750,856A
FILING DATE: 30-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94/07933
FILING DATE: 28-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 917-052-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-750-856A-16

Query Match
Best Local Similarity 4.5%; Score 6; DB 2; Length 102;
100.0%; Pred. No. 90;
Matches 6: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Oy 104 RTEVAQ 109
|||||

Db 63 RTEVAQ 68

RESULT 14
US-09-107-858-23
Sequence 23, Application US/09107858
Patent No. 6162900
GENERAL INFORMATION:
APPLICANT: Guerinet, Mary Lou et al.
TITLE OF INVENTION: METAL-REGULATED TRANSPORTERS AND USES THEREFOR
FILE REFERENCE: DCI-099CPDY
CURRENT APPLICATION NUMBER: US/09/107,858
CURRENT FILING DATE: 1998-06-30
EARLIER APPLICATION NUMBER: 08/758,621
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 23
LENGTH: 120
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-107-858-23

Query Match
Best Local Similarity 4.5%; Score 6; DB 4; Length 120;
100.0%; Pred. No. 1e+02;
Matches 6: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Oy 35 LTTSLI 40
|||||

Db 45 LTTSLI 50

RESULT 15
5169835-8
Patent No. 5169835
APPLICANT: MAI-YEE, CHAN
TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS
NUMBER OF SEQUENCES: 48
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/390,409
FILING DATE: 07-AUG-1989
SEQ ID NO: 8:
LENGTH: 144
5169835-8

Query Match
Best Local Similarity 4.5%; Score 6; DB 6; Length 144;
100.0%; Pred. No. 1.2e+02;
Matches 6: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Oy 87 EAVSLT 92
|||||

Db 69 EAVSLT 74

RESULT 16
US-09-228-986-89
Sequence 89, Application US/09228986
Patent No. 6359198
GENERAL INFORMATION:
APPLICANT: Strabala, Timothy
APPLICANT: Nieuwenhuizen, Niels
TITLE OF INVENTION: Compositions Isolated from Plant Cells

;; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
;; FILE REFERENCE: 11000/1020
;; CURRENT APPLICATION NUMBER: US/09/228,986
;; CURRENT FILING DATE: 1999-01-12
;; NUMBER OF SEQ ID NOS: 130
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 89
;; LENGTH: 154
;; TYPE: PRT
;; ORGANISM: Eucalyptus grandis
US-09-228-986-89

Query Match
Best Local Similarity 100.0%; Score 6; DB 4; Length 154;
Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 STRELT 95
Db 114 STRELT 119

RESULT 17
US-07-847-010-3
; Sequence 3, Application US/07847010
; Patent No. 5693495
; GENERAL INFORMATION:
; APPLICANT: Breiteneder, Helmo
; APPLICANT: Reikerstorfer, Arnold
; APPLICANT: Valenta, Rudolf
; APPLICANT: Hoffmann - Sommergruber, Karin
; APPLICANT: Breitenbach, Michael
; APPLICANT: Kraft, Dietrich
; APPLICANT: Rumpold, Helmut
; APPLICANT: Scheiner, Otto
; APPLICANT: Ebner, Christof
; APPLICANT: Ferreira, Fatima
; TITLE OF INVENTION: Allergens of Alder Pollen and
; TITLE OF INVENTION: Applications thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Penne & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/847,010
; FILING DATE: 01-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jones III, Harry C
; REGISTRATION NUMBER: 20,280
; REFERENCE/DOCKET NUMBER: 6530-010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 160 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Alder (Alnus sp.)
US-07-847-010-3

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 160;
Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 APEAVS 90
Db 35 APEAVS 40

RESULT 18
US-09-615-192A-388
; Sequence 388, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c4u
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 388
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-615-192A-388

Query Match
Best Local Similarity 100.0%; Score 6; DB 4; Length 161;
Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 TLMTAF 15
Db 14 TLMTAF 19

RESULT 19
US-09-370-838-125
; Sequence 125, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Koadoh
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475c1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 125
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-370-838-125

Query Match
Best Local Similarity 100.0%; Score 6; DB 4; Length 195;
Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 91 LTELTT 96
|||||
Db 83 LTELTT 88

RESULT 20
5169835-13
; Patent No. 5169835
; APPLICANT: WAI-YEE, CHAN
; TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS
; NUMBER OF SEQUENCES: 48
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/390,409
; FILING DATE: 07-AUG-1989
; SEQ ID NO:13
; LENGTH: 230
5169835-13

Query Match 4.5%; Score 6; DB 6; Length 230;
Best Local Similarity 100.0%; Pred. No.1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 87 EAVSLT 92
|||||
Db 19 EAVSLT 24

RESULT 21
US-08-318-947A-21
; Sequence 21, Application US/08318947A
; Patent No. 5798245
; GENERAL INFORMATION:
; APPLICANT: Anderson, Paul J.
; TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
; TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrie, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, NW Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,947A
FILING DATE: 06-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/133,530
FILING DATE: 07-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mack, Susan J.
REGISTRATION NUMBER: 30,951
REFERENCE/DOCKET NUMBER: A6462
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-2920
TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 244 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-318-947A-21

Query Match 4.5%; Score 6; DB 1; Length 244;

Best Local Similarity 100.0%; Pred. No.2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 91 LTELTT 96
|||||
Db 188 LTELTT 193

RESULT 22
US-08-795-303-21
; Sequence 21, Application US/08795303
; Patent No. 5948656
; GENERAL INFORMATION:
; APPLICANT: Anderson, Paul J.
; TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
; TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrie, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, NW Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,303
FILING DATE: 04-FEB-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/318,947
FILING DATE: 06-OCT-1994
APPLICATION NUMBER: 08/133,530
FILING DATE: 07-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mack, Susan J.
REGISTRATION NUMBER: 30,951
REFERENCE/DOCKET NUMBER: A6462
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-2920
TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 244 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-795-303-21

Query Match 4.5%; Score 6; DB 2; Length 244;
Best Local Similarity 100.0%; Pred. No.2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 91 LTELTT 96
|||||
Db 188 LTELTT 193

RESULT 23
US-07-857-224B-49
; Sequence 49, Application US/07857224B
; Patent No. 5958784
; GENERAL INFORMATION:
; APPLICANT: Benner, Steven A.
; TITLE OF INVENTION: Predicting Folded Structures of Proteins
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Steven A. Benner
STREET: Hadlaubstrasse 151
CITY: Zurich
STATE: none
COUNTRY: Switzerland
ZIP: (note: this is an international post code) CH-8092
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/857,224B
FILING DATE: 03/25/92
CLASSIFICATION: 436
PRIOR APPLICATION DATA: none
TELECOMMUNICATION INFORMATION:
TELEPHONE: (International) 41 1 632 2830
TELEFAX: (International) 41 1 262 2437
TELEX: none
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 259
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: protein
ORIGINAL SOURCE: human
ORGANISM: human
FEATURE: Protein kinase; Table 8 Column 56
PUBLICATION INFORMATION:
AUTHORS:
AUTHORS: Hanks, S. K.
AUTHORS: Quinn, A. M.
AUTHORS: Hunter, T.
TITLE: The protein kinase family
JOURNAL: Science
VOLUME: 241
PAGES: 42-52
DATE: 1988
US-07-857-224B-49

Query Match
Best Local Similarity 4.5%; Score 6; DB 2; Length 259;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 LTELTT 96
|||||
DB 188 LTELTT 193

RESULT 24
US-08-467-265-2
Sequence 2, Application US/08467265
Patent No. 5985612
GENERAL INFORMATION:
APPLICANT: NI, Jian
APPLICANT: Yu, Guo-Liang
APPLICANT: Gentz, Reiner
APPLICANT: Rosen, Craig A.
TITLE OF INVENTION: NATURAL KILLER CELL ENHANCED FACTOR-C
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: New Jersey
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,265
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-456
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-467-265-2

Query Match
Best Local Similarity 4.5%; Score 6; DB 2; Length 271;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 SKAKIS 74
|||||
DB 77 SKAKIS 82

RESULT 25
US-08-467-265-2
Sequence 2, Application US/08467265
Patent No. 6255079
GENERAL INFORMATION:
APPLICANT: NI, Jian
APPLICANT: Yu, Guo-Liang
APPLICANT: Gentz, Reiner
APPLICANT: Rosen, Craig A.
TITLE OF INVENTION: NATURAL KILLER CELL ENHANCED FACTOR-C
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: New Jersey
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,265
FILING DATE: 06-Jun-1995
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-456
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-467-265-2

Query Match 4.5%; Score 6; DB 4; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 69 SKAKIS 74
Db 77 SKAKIS 82

RESULT 26
US-09-407-891-2
; Sequence 2, Application US/09407891
; Patent No. 6294164
; GENERAL INFORMATION:
; APPLICANT: NI, Jian
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Gentz, Reiner
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: NATURAL KILLER CELL ENHANCED FACTOR-C
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/407,891
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/467,265
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-456
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-407-891-2

Query Match 4.5%; Score 6; DB 4; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 69 SKAKIS 74
Db 77 SKAKIS 82

RESULT 27
US-08-701-191A-32
; Sequence 32, Application US/08701191A
; Patent No. 5942428
; GENERAL INFORMATION:
; APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
; APPLICANT: and Steven R. Hubbard
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN

; TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FASTSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,191A
; FILING DATE: August 21, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wardburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/088
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-701-191A-32

Query Match 4.5%; Score 6; DB 2; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 91 LTELTT 96
Db 199 LTELTT 204

RESULT 28
US-09-090-793-3
; Sequence 3, Application US/09090793
; Patent No. 6140486
; GENERAL INFORMATION:
; APPLICANT: Calgene, LLC
; TITLE OF INVENTION: Production of polyunsaturated fatty acids by expression
; FILE REFERENCE: CGNE.131.01US
; CURRENT APPLICATION NUMBER: US/09/090,793
; CURRENT FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,650
; EARLIER FILING DATE: 1997-06-04
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Shewanella putrefaciens
US-09-090-793-3

Query Match 4.5%; Score 6; DB 4; Length 277;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 27 NPRNSS 32
Db 68 NPRNSS 73

RESULT 29

US-08-241-465B-21
; Sequence 21, Application US/08241465B
; Patent No. 5719125
; GENERAL INFORMATION:

APPLICANT: Fujio SUZUKI
APPLICANT: YUJI HIRAKI
APPLICANT: Kazuhiko TAKAHASHI
APPLICANT: Junko SUZUKI
APPLICANT: Jun KONDO
APPLICANT: Atsuko KOHARA
APPLICANT: AKIKO MORI
APPLICANT: EI YAMADA
TITLE OF INVENTION: HUMAN CHONDROMODULIN-I PROTEIN
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
COUNTRY: D.C.
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/241,465B
FILING DATE: May 11, 1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-8850
TELEFAX:
TELEX:

INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 296 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-241-465B-21

Query Match 4.5%; Score 6; DB 1; Length 296;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 128 IVPTTT 133
Db 182 IVPTTT 187

RESULT 30

US-09-173-300-26
; Sequence 26, Application US/09173300
; Patent No. 6451581
; GENERAL INFORMATION:

APPLICANT: Falco, Saverio Carl
APPLICANT: Hiltz, William D.
APPLICANT: Kinney, Anthony J.
APPLICANT: Cahoon, Rebecca E.
APPLICANT: Rafalski, J. Antoni
TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMES
FILE REFERENCE: BB-1126

CURRENT APPLICATION NUMBER: US/09/173,300
CURRENT FILING DATE: 1998-10-15
EARLIER APPLICATION NUMBER: 60/063,423
EARLIER FILING DATE: 1997 October 28
NUMBER OF SEQ ID NOS: 54
SOFTWARE: Microsoft Word Version 7.0A
SEQ ID NO 26
LENGTH: 297
TYPE: PRT
ORGANISM: Oryza sativa
US-09-173-300-26

Query Match 4.5%; Score 6; DB 4; Length 297;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 69 SKAKIS 74
Db 57 SKAKIS 62

RESULT 31

US-08-946-528-7
; Sequence 7, Application US/08946528
; Patent No. 5958746
; GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Nell C.
TITLE OF INVENTION: ELECTRON TRANSPORT PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/946,528
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0406 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 300 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 162941
US-08-946-528-7

Query Match 4.5%; Score 6; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 DKAPEA 88
Db 239 DKAPEA 244

RESULT 32

US-09-173-300-15
; Sequence 15, Application US/09173300
; Patent No. 6451581
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Hiltz, William D.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB-1126
; CURRENT APPLICATION NUMBER: US/09/173.300
; EARLIER FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: 60/063,423
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 15
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-173-300-15

Query Match

Best Local Similarity 4.5%; Score 6; DB 4; Length 307;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 HAVKOT 47
Db 72 HAVKOT 77

RESULT 33

US-09-135-639-4
; Sequence 4, Application US/09135639
; Patent No. 6004793
; GENERAL INFORMATION:
; APPLICANT: LUNNEN, KEITH D.
; APPLICANT: DALTON MICHAEL A.,
; APPLICANT: WILSON, GEOFREY G.
; APPLICANT: XU, SHUANG-YONG
; TITLE OF INVENTION: Method For Cloning And Producing The Avari Restriction
; TITLE OF INVENTION: Endonuclease In E. coli And Purification Of The
; FILE REFERENCE: Avari
; CURRENT APPLICATION NUMBER: US/09/135,639
; CURRENT FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Anabaena variabilis
US-09-135-639-4

Query Match

Best Local Similarity 4.5%; Score 6; DB 3; Length 315;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 RHSLKP 118
Db 262 RHSLKP 267

RESULT 34

US-08-241-465B-19
; Sequence 19, Application US/08241465B

; Patent No. 5719125
; GENERAL INFORMATION:

; APPLICANT: Fujio SUZUKI
; APPLICANT: Kazuhiro TAKAHASHI
; APPLICANT: Junko SUZUKI
; APPLICANT: Jun Kondo
; APPLICANT: Atsuko KOHARA
; APPLICANT: AKIKO MORI
; APPLICANT: EI YAMADA
; TITLE OF INVENTION: HUMAN CHONDROMODULIN-I PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Menderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; COUNTRY: D.C.
; ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/241,465B
FILING DATE: May 11, 1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-8850
TELEFAX:

TELEX:

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:
LENGTH: 334 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-241-465B-19

Query Match

Best Local Similarity 4.5%; Score 6; DB 1; Length 334;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 IVPPTT 133
Db 220 IVPPTT 225

RESULT 35

US-08-241-465B-20
; Sequence 20, Application US/08241465B
; Patent No. 5719125
; GENERAL INFORMATION:
; APPLICANT: Fujio SUZUKI
; APPLICANT: Kazuhiro TAKAHASHI
; APPLICANT: Junko SUZUKI
; APPLICANT: Jun Kondo
; APPLICANT: Atsuko KOHARA
; APPLICANT: AKIKO MORI
; APPLICANT: EI YAMADA
; TITLE OF INVENTION: HUMAN CHONDROMODULIN-I PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Menderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; COUNTRY: D.C.
; ZIP: 20005

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/241,465B
FILING DATE: May 11, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mairen M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 334 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-241-465B-20

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 334;
Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 IVPPTT 133
|||||
DB 220 IVPPTT 225

RESULT 36
US-08-758-621-14
; Sequence 14, Application US/08758621
; Patent No. 5846821
; GENERAL INFORMATION:
; APPLICANT: Gueriot, Mary Lou, and Elide, David J.
; TITLE OF INVENTION: Metal-Regulated Transporters and Uses Therefor
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHYE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/758,621
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/018,578
; FILING DATE: 29-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Silverl, Jean M.
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: DCI-099CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 345 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
US-08-758-621-14

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 345;
Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 LTTSLI 40
|||||
DB 50 LTTSLI 55

RESULT 37
US-09-107-858-14
; Sequence 14, Application US/09107858
; Patent No. 6162900
; GENERAL INFORMATION:
; APPLICANT: Gueriot, Mary Lou et al.
; TITLE OF INVENTION: METAL-REGULATED TRANSPORTERS AND USES THEREFOR
; FILE REFERENCE: DCI-099CPDV
; CURRENT APPLICATION NUMBER: US/09/107,858
; CURRENT FILING DATE: 1998-06-30
; EARLIER APPLICATION NUMBER: 08/758,621
; EARLIER FILING DATE: 1996-11-27
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-107-858-14

Query Match
Best Local Similarity 100.0%; Score 6; DB 4; Length 345;
Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 LTTSLI 40
|||||
DB 50 LTTSLI 55

RESULT 38
US-09-134-001C-5513
; Sequence 5513, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5513
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5513

Query Match
Best Local Similarity 100.0%; Score 6; DB 4; Length 348;
Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 SUTELT 95
|||||
DB 253 SUTELT 258

RESULT 39
US-08-202-056-1
```

```
Sequence 1, Application US/08202056
Patent No. 5440021
GENERAL INFORMATION:
APPLICANT: Chuntcharapal, Anan
APPLICANT: Hebert, Caroline
APPLICANT: Kim, Kyung Jin
APPLICANT: Lee, James
TITLE OF INVENTION: Antibodies to Human IL-8 Type B Receptor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patln (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/202,056
FILING DATE: 25-FEB-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/677211
FILING DATE: 29-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: 706P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-202-056-1

Query Match
Best Local Similarity 4.5%; Score 6; DB 1; Length 350;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 118 PCMLET 123
Db 29 PCMLET 34

RESULT 40
US-08-076-093A-2
Sequence 2, Application US/08076093A
Patent No. 5543503
GENERAL INFORMATION:
APPLICANT: Chuntcharapal, Anan
APPLICANT: Lee, James
APPLICANT: Hebert, Caroline
APPLICANT: Jin Kim, K.
TITLE OF INVENTION: Antibodies to Human PFAA Receptors
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatln (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/076,093A
FILING DATE: 11-Jun-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/810782
FILING DATE: 19-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/677211
FILING DATE: 29-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: 706P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: Amino Acid
TOPOLOGY: linear
US-08-076-093A-2

Query Match
Best Local Similarity 4.5%; Score 6; DB 1; Length 350;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 118 PCMLET 123
Db 29 PCMLET 34

Search completed: April 28, 2003, 16:15:19
Job time : 33 secs
```

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 16:15:08 ; Search time 45 Seconds
(Without alignments)
238.609 Million cell updates/sec

Title: US-10-048-196-2

Perfect score: 134
Sequence: 1 MKKILYVATMTAFATLASC.....SLKPCMLETVNAFVPTTR 134

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 301932 seqs, 80129803 residues

Word size : 0

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/PCR_US07_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	5.2	345	10 US-09-815-242-11068	Sequence 11068, A
2	7	5.2	596	10 US-09-862-027-21	Sequence 21, App1
3	6	4.5	26	9 US-10-097-065-562	Sequence 562, App
4	6	4.5	29	10 US-09-864-761-48009	Sequence 48009, A
5	6	4.5	42	10 US-09-864-761-35812	Sequence 35812, A
6	6	4.5	70	9 US-10-097-065-561	Sequence 561, App
7	6	4.5	76	10 US-09-815-242-4896	Sequence 4896, Ap
8	6	4.5	77	10 US-09-864-761-41016	Sequence 41016, A
9	6	4.5	77	10 US-09-864-761-42371	Sequence 42371, A
10	6	4.5	81	10 US-09-815-242-10863	Sequence 10863, A
11	6	4.5	88	9 US-10-007-280A-238	Sequence 238, App
12	6	4.5	116	9 US-10-101-464A-499	Sequence 499, App
13	6	4.5	131	10 US-09-895-728-6	Sequence 728, App
14	6	4.5	140	9 US-10-101-464A-738	Sequence 738, App
15	6	4.5	153	10 US-09-764-864-1364	Sequence 1364, Ap
16	6	4.5	154	9 US-10-101-464A-89	Sequence 89, App1
17	6	4.5	195	9 US-09-854-133-125	Sequence 125, App
18	6	4.5	195	10 US-09-738-973-125	Sequence 125, App
19	6	4.5	210	9 US-10-072-349-90	Sequence 90, App1

20	6	4.5	210	10 US-09-764-855-90	Sequence 90, App1
21	6	4.5	211	9 US-09-738-626-5266	Sequence 5266, Ap
22	6	4.5	261	9 US-09-738-626-5335	Sequence 6335, Ap
23	6	4.5	271	10 US-09-911-346-2	Sequence 2, App1
24	6	4.5	276	10 US-09-764-864-925	Sequence 925, App
25	6	4.5	287	9 US-09-738-626-4972	Sequence 4972, Ap
26	6	4.5	297	12 US-10-027-450-26	Sequence 26, App1
27	6	4.5	299	9 US-09-510-332-9	Sequence 9, App1
28	6	4.5	309	10 US-09-393-634-43	Sequence 43, App1
29	6	4.5	307	12 US-10-027-450-15	Sequence 15, App1
30	6	4.5	309	10 US-09-801-368-286	Sequence 286, App
31	6	4.5	311	9 US-10-260-877-32	Sequence 32, App1
32	6	4.5	332	9 US-09-964-899-7	Sequence 7, App1
33	6	4.5	336	10 US-09-815-242-5618	Sequence 5618, Ap
34	6	4.5	336	10 US-09-815-242-12540	Sequence 12540, A
35	6	4.5	343	9 US-10-116-016-37	Sequence 37, App1
36	6	4.5	343	10 US-09-764-848-37	Sequence 37, App1
37	6	4.5	350	10 US-09-104-063-2	Sequence 2, App1
38	6	4.5	350	10 US-09-782-980-83	Sequence 83, App1
39	6	4.5	350	10 US-09-884-430-7	Sequence 7, App1
40	6	4.5	364	10 US-09-992-807-3	Sequence 3, App1
41	6	4.5	366	9 US-10-143-849-6	Sequence 6, App1
42	6	4.5	375	10 US-09-815-242-12383	Sequence 12383, A
43	6	4.5	376	10 US-09-815-242-5453	Sequence 5453, Ap
44	6	4.5	378	9 US-10-108-605-123	Sequence 123, App
45	6	4.5	380	9 US-09-165-522-16	Sequence 16, App1
46	6	4.5	381	10 US-09-815-242-10962	Sequence 10962, A
47	6	4.5	388	9 US-10-125-792-6	Sequence 6, App1
48	6	4.5	412	10 US-09-864-761-37085	Sequence 37085, A
49	6	4.5	418	10 US-09-816-664-2	Sequence 2, App1
50	6	4.5	420	10 US-09-729-674-172	Sequence 172, App
51	6	4.5	431	9 US-09-286-488-34	Sequence 34, App1
52	6	4.5	431	10 US-09-737-778-34	Sequence 34, App1
53	6	4.5	440	9 US-10-011-588-31	Sequence 31, App1
54	6	4.5	440	10 US-09-815-242-10475	Sequence 10475, A
55	6	4.5	441	9 US-10-011-588-11	Sequence 11, App1
56	6	4.5	454	10 US-09-866-582-8	Sequence 8, App1
57	6	4.5	497	10 US-09-846-5908-10	Sequence 10, App1
58	6	4.5	501	10 US-09-951-622-9	Sequence 9, App1
59	6	4.5	526	9 US-09-738-626-5160	Sequence 5160, Ap
60	6	4.5	536	9 US-09-977-260-13	Sequence 13, App1
61	6	4.5	536	9 US-09-929-266-10	Sequence 10, App1
62	6	4.5	536	9 US-09-977-261-13	Sequence 13, App1
63	6	4.5	536	10 US-09-977-269-13	Sequence 13, App1
64	6	4.5	537	9 US-10-047-542-74	Sequence 74, App1
65	6	4.5	562	9 US-09-992-238-14	Sequence 14, App1
66	6	4.5	572	9 US-10-185-991-2	Sequence 2, App1
67	6	4.5	572	9 US-09-992-238-13	Sequence 13, App1
68	6	4.5	585	10 US-09-925-297-800	Sequence 800, App
69	6	4.5	626	9 US-09-738-626-5957	Sequence 6957, Ap
70	6	4.5	648	9 US-09-975-719-221	Sequence 221, App
71	6	4.5	660	9 US-10-101-464A-808	Sequence 808, App
72	6	4.5	727	10 US-09-923-435A-2	Sequence 2, App1
73	6	4.5	774	10 US-09-833-435A-8	Sequence 8, App1
74	6	4.5	824	9 US-10-011-588-33	Sequence 33, App1
75	6	4.5	829	9 US-10-055-364-46	Sequence 46, App1
76	6	4.5	833	9 US-09-470-276-54	Sequence 54, App1
77	6	4.5	896	9 US-10-217-096-6	Sequence 6, App1
78	6	4.5	941	9 US-10-125-792-8	Sequence 8, App1
79	6	4.5	941	9 US-10-125-792-10	Sequence 10, App1
80	6	4.5	954	9 US-10-101-464A-911	Sequence 911, App
81	6	4.5	954	10 US-09-815-242-11338	Sequence 11338, A
82	6	4.5	1013	10 US-09-942-366-5	Sequence 5, App1
83	6	4.5	1013	10 US-09-893-519A-14	Sequence 20, App1
84	6	4.5	1023	9 US-09-893-519A-14	Sequence 14, App1
85	6	4.5	1069	9 US-10-007-270-24	Sequence 24, App1
86	6	4.5	2025	10 US-09-815-242-5703	Sequence 5703, Ap
87	6	4.5	2437	10 US-09-815-242-5844	Sequence 5834, Ap
88	6	4.5	3158	10 US-09-815-242-12611	Sequence 12611, A
89	6	4.5	6281	10 US-09-815-242-12996	Sequence 12996, A
90	6	3.7	6	10 US-09-911-838-176	Sequence 176, App
91	5	3.7	6	10 US-09-911-838-178	Sequence 178, App
92	5	3.7	7	9 US-09-996-288-164	Sequence 164, App

93	3.7	7	9	US-09-990-186-847	Sequence 847, App	166	3.7	35	9	US-09-746-660A-100	Sequence 100, App
94	3.7	7	9	US-09-990-186-848	Sequence 848, App	167	3.7	35	10	US-09-810-310-1	Sequence 1, App1
95	3.7	7	9	US-09-990-186-892	Sequence 892, App	168	3.7	35	10	US-09-810-310-8	Sequence 8, App1
96	3.7	7	9	US-09-990-186-901	Sequence 901, App	169	3.7	36	10	US-09-864-761-47897	Sequence 47897, A
97	3.7	7	9	US-09-990-186-905	Sequence 905, App	170	3.7	37	10	US-09-864-761-477013	Sequence 477013, A
98	3.7	7	9	US-09-990-186-906	Sequence 906, App	171	3.7	38	9	US-09-789-054A-8	Sequence 8, App1
99	3.7	7	9	US-09-990-186-3780	Sequence 3780, App	172	3.7	38	9	US-10-223-047-21	Sequence 21, App1
100	3.7	7	10	US-09-989-789-847	Sequence 847, App	173	3.7	38	10	US-09-864-761-44497	Sequence 44497, A
101	3.7	7	10	US-09-989-789-848	Sequence 848, App	174	3.7	40	10	US-09-864-761-34456	Sequence 34456, A
102	3.7	7	10	US-09-989-789-892	Sequence 892, App	175	3.7	41	10	US-09-864-761-42451	Sequence 42451, A
103	3.7	7	10	US-09-989-789-901	Sequence 901, App	176	3.7	41	10	US-09-864-761-42451	Sequence 42451, A
104	3.7	7	10	US-09-989-789-905	Sequence 905, App	177	3.7	43	9	US-09-948-820-71	Sequence 71, App1
105	3.7	7	10	US-09-989-789-906	Sequence 906, App	178	3.7	43	10	US-09-864-761-34994	Sequence 34994, A
106	3.7	7	10	US-09-989-789-906	Sequence 906, App	179	3.7	43	10	US-09-864-761-43463	Sequence 43463, A
107	3.7	7	10	US-09-989-789-3780	Sequence 3780, App	180	3.7	45	8	US-08-424-550B-199	Sequence 199, App
108	3.7	7	10	US-09-911-838-117	Sequence 117, App	181	3.7	46	10	US-09-864-761-48128	Sequence 48128, A
109	3.7	7	10	US-09-911-838-117	Sequence 117, App	182	3.7	46	10	US-09-864-761-48604	Sequence 48604, A
110	3.7	7	9	US-10-007-363-1	Sequence 1, App1	183	3.7	47	9	US-09-984-271-122	Sequence 122, App
111	3.7	8	10	US-09-756-283A-80	Sequence 80, App1	184	3.7	56	10	US-09-764-877-1100	Sequence 1100, App
112	3.7	8	12	US-10-042-203-38	Sequence 38, App1	185	3.7	57	10	US-09-983-802-534	Sequence 534, App
113	3.7	8	12	US-10-007-761-1	Sequence 1, App1	186	3.7	57	10	US-09-983-802-534	Sequence 33897, A
114	3.7	9	10	US-09-779-308-117	Sequence 117, App	187	3.7	58	9	US-09-963-802-203	Sequence 203, App
115	3.7	9	10	US-09-779-308-117	Sequence 117, App	188	3.7	58	10	US-09-864-761-43110	Sequence 43110, A
116	3.7	9	10	US-09-779-308-404	Sequence 404, App	189	3.7	58	10	US-09-864-761-43623	Sequence 43623, A
117	3.7	9	10	US-09-779-308-487	Sequence 487, App	190	3.7	59	9	US-10-044-359-6	Sequence 6, App1
118	3.7	9	10	US-09-779-308-487	Sequence 487, App	191	3.7	59	9	US-09-764-904-46	Sequence 46, App1
119	3.7	9	10	US-09-779-308-508	Sequence 508, App	192	3.7	59	10	US-09-825-302-658	Sequence 658, App
120	3.7	9	10	US-09-779-308-581	Sequence 581, App	193	3.7	59	10	US-09-864-761-40483	Sequence 40483, A
121	3.7	9	10	US-09-779-308-598	Sequence 598, App	194	3.7	59	10	US-09-764-860-311	Sequence 311, App
122	3.7	9	10	US-09-779-308-702	Sequence 702, App	195	3.7	60	9	US-10-002-344A-269	Sequence 269, App
123	3.7	10	10	US-09-779-308-119	Sequence 119, App	196	3.7	60	10	US-09-864-761-38078	Sequence 38078, A
124	3.7	10	10	US-09-779-308-147	Sequence 147, App	197	3.7	60	10	US-09-864-761-45784	Sequence 45784, A
125	3.7	10	10	US-09-779-308-426	Sequence 426, App	198	3.7	60	10	US-09-864-761-45784	Sequence 45784, A
126	3.7	10	10	US-09-779-308-434	Sequence 434, App	199	3.7	62	9	US-09-815-242-4871	Sequence 4871, App
127	3.7	10	10	US-09-779-308-551	Sequence 551, App	200	3.7	62	9	US-09-776-124A-148	Sequence 148, App
128	3.7	11	1	US-08-779-457-43	Sequence 43, App1	201	3.7	62	10	US-09-864-761-37032	Sequence 37032, A
129	3.7	11	1	US-10-214-802-43	Sequence 43, App1	202	3.7	64	9	US-09-796-692-1257	Sequence 1257, App
130	3.7	11	9	US-10-033-350-7	Sequence 7, App1	203	3.7	65	10	US-09-864-761-43662	Sequence 43662, App
131	3.7	11	9	US-09-056-160B-17	Sequence 17, App1	204	3.7	68	9	US-09-925-299-1472	Sequence 1472, App
132	3.7	11	10	US-09-825-584-1	Sequence 1, App1	205	3.7	68	10	US-09-864-761-46225	Sequence 46225, A
133	3.7	11	10	US-09-811-384-8	Sequence 8, App1	206	3.7	68	10	US-09-925-299-1472	Sequence 1472, App
134	3.7	12	10	US-09-911-838-208	Sequence 208, App	207	3.7	69	9	US-09-764-877-1195	Sequence 1195, App
135	3.7	13	10	US-09-459-749D-5	Sequence 5, App1	208	3.7	69	9	US-10-200-910-14	Sequence 14, App1
136	3.7	13	10	US-09-985-065-1	Sequence 1, App1	209	3.7	70	10	US-09-864-761-41178	Sequence 41178, A
137	3.7	13	10	US-09-911-838-5	Sequence 5, App1	210	3.7	70	10	US-09-815-242-10523	Sequence 10523, A
138	3.7	13	10	US-09-911-838-5	Sequence 5, App1	211	3.7	72	10	US-09-764-860-335	Sequence 335, App
139	3.7	14	10	US-09-826-177-53	Sequence 53, App1	212	3.7	72	10	US-09-864-761-47147	Sequence 47147, A
140	3.7	14	10	US-09-826-177-57	Sequence 57, App1	213	3.7	73	9	US-10-083-357-1213	Sequence 1213, App
141	3.7	15	10	US-09-823-153-9	Sequence 9, App1	214	3.7	74	9	US-09-864-761-43084	Sequence 43084, A
142	3.7	15	10	US-09-823-153-9	Sequence 9, App1	215	3.7	74	9	US-09-796-692-2275	Sequence 2275, App
143	3.7	15	10	US-09-823-153-5	Sequence 5, App1	216	3.7	74	10	US-09-864-761-35678	Sequence 35678, A
144	3.7	16	10	US-09-823-153-5	Sequence 5, App1	217	3.7	74	10	US-09-864-761-43326	Sequence 43326, A
145	3.7	16	10	US-09-823-153-5	Sequence 5, App1	218	3.7	75	10	US-09-925-297-821	Sequence 821, App
146	3.7	18	9	US-09-885-065-3	Sequence 3, App1	219	3.7	75	9	US-09-272-975-6	Sequence 6, App1
147	3.7	19	9	US-10-062-710-37	Sequence 37, App1	220	3.7	77	9	US-09-746-783-138	Sequence 138, App
148	3.7	21	10	US-09-880-748-2885	Sequence 2885, App	221	3.7	77	9	US-09-864-761-37029	Sequence 37029, A
149	3.7	21	10	US-09-864-761-37029	Sequence 37029, A	222	3.7	77	12	US-10-001-870-163	Sequence 163, App
150	3.7	23	9	US-09-969-730-202	Sequence 202, App	223	3.7	80	9	US-10-091-572-230	Sequence 230, App
151	3.7	24	9	US-09-774-639-198	Sequence 198, App	224	3.7	83	9	US-09-864-761-36212	Sequence 36212, A
152	3.7	25	9	US-09-973-025-55	Sequence 55, App1	225	3.7	83	10	US-09-160-116-4	Sequence 4, App1
153	3.7	25	9	US-09-899-303-55	Sequence 55, App1	226	3.7	84	10	US-09-864-761-37787	Sequence 37787, A
154	3.7	25	10	US-09-776-724A-263	Sequence 263, App	227	3.7	84	10	US-09-864-761-43315	Sequence 43315, A
155	3.7	25	10	US-09-864-761-39706	Sequence 39706, App	228	3.7	84	10	US-09-754-997A-324	Sequence 24, App1
156	3.7	26	10	US-09-864-761-41414	Sequence 41414, A	229	3.7	84	10	US-09-864-761-43315	Sequence 330, App
157	3.7	26	10	US-09-864-761-45975	Sequence 45975, A	230	3.7	86	9	US-09-925-299-948	Sequence 948, App
158	3.7	27	10	US-09-864-761-48309	Sequence 48309, A	231	3.7	87	9	US-09-825-302-658	Sequence 45253, A
159	3.7	29	9	US-10-011-585A-215	Sequence 215, App	232	3.7	87	10	US-09-864-761-45835	Sequence 948, App
160	3.7	30	10	US-09-864-761-45724	Sequence 45724, A	233	3.7	87	10	US-09-925-299-948	Sequence 948, App
161	3.7	31	10	US-09-864-761-48944	Sequence 48944, A	234	3.7	87	10	US-09-864-761-45835	Sequence 948, App
162	3.7	31	10	US-09-938-315-106	Sequence 106, App	235	3.7	88	9	US-09-774-639-116	Sequence 116, App
163	3.7	33	10	US-09-864-761-41041	Sequence 41041, A	236	3.7	89	9	US-10-007-280A-217	Sequence 217, App
164	3.7	34	10	US-09-938-719-13	Sequence 13, App1	237	3.7	90	10	US-09-920-920-75	Sequence 75, App1
165	3.7	34	10	US-09-938-719-13	Sequence 13, App1	238	3.7	91	10	US-09-864-761-43070	Sequence 43070, A
166	3.7	34	10	US-09-938-719-13	Sequence 13, App1	239	3.7	92	9	US-09-966-480-241	Sequence 241, App

239	5	3.7	92	9	US-10-092-154-818	Sequence 818, App	312	5	3.7	116	10	US-09-220-920-4	Sequence 4, Appl1
240	5	3.7	92	12	US-09-764-847-818	Sequence 818, App	313	5	3.7	116	10	US-09-220-920-35	Sequence 35, Appl1
241	5	3.7	92	10	US-10-078-929-12	Sequence 12, App	314	5	3.7	116	10	US-09-804-615-11	Sequence 11, Appl1
242	5	3.7	93	9	US-09-783-931-44	Sequence 44, Appl1	315	5	3.7	116	10	US-09-925-300-1028	Sequence 1028, App
243	5	3.7	93	10	US-09-908-322-44	Sequence 44, Appl1	316	5	3.7	117	10	US-09-764-887-194	Sequence 194, App
244	5	3.7	94	10	US-09-815-242-11371	Sequence 11371, A	317	5	3.7	118	9	US-09-430-029-8	Sequence 8, Appl1
245	5	3.7	94	10	US-09-764-864-1035	Sequence 1035, Ap	318	5	3.7	118	10	US-09-867-550-1670	Sequence 1670, Ap
246	5	3.7	96	10	US-09-220-920-19	Sequence 19, Appl1	319	5	3.7	118	10	US-09-925-301-1643	Sequence 1643, Ap
247	5	3.7	96	10	US-09-220-920-33	Sequence 33, Appl1	320	5	3.7	119	10	US-09-925-301-1361	Sequence 1361, Ap
248	5	3.7	96	10	US-09-263-959-1192	Sequence 1192, Ap	321	5	3.7	120	9	US-09-809-391-926	Sequence 34, Appl1
249	5	3.7	97	9	US-10-101-664A-651	Sequence 651, App	322	5	3.7	120	9	US-09-609-929-126	Sequence 46, App
250	5	3.7	97	9	US-10-081-572-358	Sequence 358, App	323	5	3.7	120	9	US-10-043-487-343	Sequence 343, App
251	5	3.7	98	9	US-09-925-299-1067	Sequence 1067, App	324	5	3.7	120	10	US-09-864-761-44932	Sequence 44932, A
252	5	3.7	98	10	US-09-925-299-1067	Sequence 1067, Ap	325	5	3.7	121	9	US-09-738-626-5331	Sequence 5331, App
253	5	3.7	98	10	US-09-098-079-21	Sequence 21, Appl1	326	5	3.7	121	9	US-09-866-605A-688	Sequence 688, App
254	5	3.7	99	9	US-09-764-868-1055	Sequence 1055, App	327	5	3.7	122	9	US-09-738-626-5556	Sequence 5556, Ap
255	5	3.7	99	9	US-09-764-872-327	Sequence 327, App	328	5	3.7	122	10	US-09-925-297-721	Sequence 721, App
256	5	3.7	100	9	US-10-032-162-2	Sequence 2, Appl1	329	5	3.7	123	9	US-09-969-384-21	Sequence 21, Appl1
257	5	3.7	100	10	US-09-811-384-7	Sequence 7, Appl1	330	5	3.7	123	9	US-10-043-487-333	Sequence 333, App
258	5	3.7	100	10	US-09-925-297-722	Sequence 722, App	331	5	3.7	124	9	US-09-764-868-1196	Sequence 1196, Ap
259	5	3.7	101	9	US-09-796-692-821	Sequence 821, App	332	5	3.7	124	9	US-09-975-719-53	Sequence 53, Appl1
260	5	3.7	101	9	US-09-796-692-930	Sequence 930, App	333	5	3.7	124	12	US-10-052-545-12	Sequence 12, Appl1
261	5	3.7	101	9	US-09-796-692-934	Sequence 934, App	334	5	3.7	125	9	US-09-738-626-5392	Sequence 5392, Ap
262	5	3.7	101	9	US-09-796-692-973	Sequence 973, App	335	5	3.7	125	9	US-09-738-626-5499	Sequence 5499, App
263	5	3.7	101	9	US-09-796-692-984	Sequence 984, App	336	5	3.7	126	9	US-10-101-664A-100	Sequence 100, App
264	5	3.7	101	9	US-09-796-692-1313	Sequence 1313, Ap	337	5	3.7	127	10	US-09-815-242-10432	Sequence 10432, A
265	5	3.7	101	9	US-09-796-692-1319	Sequence 1319, Ap	338	5	3.7	127	10	US-09-867-550-420	Sequence 420, App
266	5	3.7	101	9	US-09-796-692-1941	Sequence 1941, Ap	339	5	3.7	130	10	US-09-864-761-45365	Sequence 45365, A
267	5	3.7	102	10	US-09-205-658-258	Sequence 258, App	340	5	3.7	130	10	US-09-925-301-1479	Sequence 1479, Ap
268	5	3.7	103	10	US-09-767-870-12	Sequence 12, Appl1	341	5	3.7	131	10	US-09-925-301-1471	Sequence 1471, Ap
269	5	3.7	103	10	US-09-867-550-856	Sequence 856, App	342	5	3.7	131	10	US-09-764-864-924	Sequence 924, App
270	5	3.7	104	9	US-10-072-349-111	Sequence 111, App	343	5	3.7	131	10	US-09-764-864-1363	Sequence 1363, App
271	5	3.7	104	9	US-10-079-854-118	Sequence 118, App	344	5	3.7	132	9	US-09-789-054A-2	Sequence 2, Appl1
272	5	3.7	104	10	US-09-764-878-118	Sequence 118, App	345	5	3.7	132	10	US-10-101-664A-572	Sequence 572, App
273	5	3.7	104	10	US-09-764-855-111	Sequence 111, App	346	5	3.7	132	10	US-09-764-864-1523	Sequence 1523, App
274	5	3.7	105	9	US-10-013-379-33	Sequence 33, Appl1	347	5	3.7	133	10	US-09-795-006A-151	Sequence 151, App
275	5	3.7	105	9	US-10-068-564-26	Sequence 26, Appl1	348	5	3.7	133	9	US-09-852-209A-11	Sequence 11, Appl1
276	5	3.7	105	10	US-10-079-854-192	Sequence 192, App	349	5	3.7	133	9	US-10-114-693-123	Sequence 123, App
277	5	3.7	105	10	US-09-764-878-1192	Sequence 192, App	350	5	3.7	134	9	US-09-986-480-366	Sequence 366, App
278	5	3.7	105	10	US-09-989-903-26	Sequence 26, Appl1	351	5	3.7	134	10	US-09-764-864-1455	Sequence 1425, App
279	5	3.7	106	10	US-09-864-761-3598	Sequence 3598, A	352	5	3.7	135	9	US-09-973-457-5	Sequence 5, Appl1
280	5	3.7	107	9	US-10-066-127-10	Sequence 10, Appl1	353	5	3.7	135	9	US-09-738-626-4481	Sequence 4481, App
281	5	3.7	107	9	US-09-796-692-1468	Sequence 1468, Ap	354	5	3.7	135	9	US-10-101-664A-563	Sequence 563, App
282	5	3.7	107	9	US-09-796-692-1910	Sequence 1910, Ap	355	5	3.7	135	10	US-09-804-615-40	Sequence 40, Appl1
283	5	3.7	107	9	US-09-796-692-1950	Sequence 1950, App	356	5	3.7	135	12	US-10-074-527-6	Sequence 6, Appl1
284	5	3.7	107	10	US-09-864-761-46046	Sequence 46046, A	357	5	3.7	136	9	US-09-738-626-4446	Sequence 4446, App
285	5	3.7	108	9	US-09-910-483-27	Sequence 27, Appl1	358	5	3.7	137	9	US-09-796-692-853	Sequence 853, App
286	5	3.7	108	9	US-09-910-483-31	Sequence 31, Appl1	359	5	3.7	137	9	US-09-848-616-140	Sequence 140, App
287	5	3.7	108	9	US-09-910-483-35	Sequence 35, Appl1	360	5	3.7	139	10	US-09-815-242-5364	Sequence 5364, App
288	5	3.7	108	9	US-09-910-483-38	Sequence 38, Appl1	361	5	3.7	140	10	US-09-920-920-5	Sequence 5, Appl1
289	5	3.7	108	9	US-09-910-483-42	Sequence 42, Appl1	362	5	3.7	140	10	US-09-864-761-46650	Sequence 46650, A
290	5	3.7	109	9	US-10-083-357-775	Sequence 775, App	363	5	3.7	140	10	US-09-804-615-10	Sequence 10, Appl1
291	5	3.7	109	9	US-10-091-504-1033	Sequence 1033, Ap	364	5	3.7	140	10	US-09-764-864-1105	Sequence 1105, Ap
292	5	3.7	109	10	US-09-764-869-1033	Sequence 1033, Ap	365	5	3.7	140	10	US-09-764-864-11525	Sequence 1525, Ap
293	5	3.7	110	9	US-09-764-876-22	Sequence 22, Appl1	366	5	3.7	141	10	US-09-764-864-1220	Sequence 1220, Ap
294	5	3.7	112	10	US-09-925-301-1481	Sequence 1481, Ap	367	5	3.7	142	10	US-09-815-242-11039	Sequence 11039, A
295	5	3.7	113	9	US-09-851-138-92	Sequence 28, Appl1	368	5	3.7	143	9	US-09-738-626-5904	Sequence 5904, App
296	5	3.7	113	9	US-10-012-140-28	Sequence 28, Appl1	369	5	3.7	143	10	US-09-907-539-2	Sequence 2, Appl1
297	5	3.7	113	10	US-09-220-920-3	Sequence 3, Appl1	370	5	3.7	144	9	US-09-929-230-8	Sequence 8, Appl1
298	5	3.7	113	10	US-09-220-920-34	Sequence 34, Appl1	371	5	3.7	144	10	US-09-220-920-36	Sequence 36, Appl1
299	5	3.7	113	10	US-09-804-615-12	Sequence 12, Appl1	372	5	3.7	145	10	US-09-775-046-5	Sequence 5, Appl1
300	5	3.7	113	10	US-09-867-550-1916	Sequence 1916, Ap	373	5	3.7	146	9	US-09-764-868-1017	Sequence 1017, Ap
301	5	3.7	113	10	US-09-925-300-1576	Sequence 1576, Ap	374	5	3.7	146	9	US-09-764-868-882	Sequence 882, App
302	5	3.7	113	12	US-10-078-929-2	Sequence 2, Appl1	375	5	3.7	147	10	US-09-864-761-36066	Sequence 36066, A
303	5	3.7	114	9	US-09-808-602-111	Sequence 111, App	376	5	3.7	148	10	US-09-864-761-36066	Sequence 36066, A
304	5	3.7	114	9	US-10-092-154-979	Sequence 979, App	377	5	3.7	149	9	US-10-013-379-9	Sequence 9, Appl1
305	5	3.7	114	10	US-09-864-761-34736	Sequence 34736, A	378	5	3.7	151	10	US-09-776-724A-1069	Sequence 262, App
306	5	3.7	114	10	US-09-804-615-37	Sequence 37, App	379	5	3.7	151	9	US-09-738-626-3689	Sequence 3689, App
307	5	3.7	114	10	US-09-731-872-394	Sequence 394, App	380	5	3.7	152	9	US-09-815-242-12642	Sequence 12642, A
308	5	3.7	114	10	US-09-764-847-979	Sequence 979, App	381	5	3.7	152	10	US-09-815-242-12879	Sequence 12879, A
309	5	3.7	114	10	US-09-263-959-330	Sequence 330, App	382	5	3.7	152	10	US-09-861-451A-66	Sequence 66, Appl1
310	5	3.7	114	10	US-09-925-300-1621	Sequence 1621, Ap	383	5	3.7	152	10	US-09-939-980-264	Sequence 264, App
311	5	3.7	116	9	US-09-895-913A-178	Sequence 178, App	384	5	3.7	152	10		

385	5	3.7	153	12	US-10-078-929-58	Sequence 58, Appl	458	5	3.7	191	10	US-09-815-242-14038	Sequence 14038, A
386	5	3.7	154	10	US-09-925-301-1478	Sequence 1478, Ap	459	5	3.7	191	10	US-09-893-737-224	Sequence 224, Ap
387	5	3.7	155	9	US-09-764-868-1186	Sequence 1186, Ap	460	5	3.7	193	9	US-09-791-279-174	Sequence 174, Ap
388	5	3.7	155	10	US-09-925-301-1561	Sequence 1561, Ap	461	5	3.7	194	9	US-09-738-626-4966	Sequence 4966, Ap
389	5	3.7	156	10	US-09-771-161A-104	Sequence 104, Ap	462	5	3.7	195	9	US-09-738-626-4581	Sequence 4581, Ap
390	5	3.7	157	9	US-09-738-626-5954	Sequence 5954, Ap	463	5	3.7	195	9	US-10-222-162-17	Sequence 17, Appl
391	5	3.7	157	9	US-09-984-271-212	Sequence 212, App	464	5	3.7	197	10	US-09-905-114-2	Sequence 2, Appl1
392	5	3.7	157	10	US-09-867-550-1340	Sequence 1340, Ap	465	5	3.7	198	10	US-09-815-242-13889	Sequence 13889, A
393	5	3.7	158	9	US-09-973-025-12	Sequence 12, Appl	466	5	3.7	199	10	US-09-815-837-56	Sequence 56, Appl
394	5	3.7	158	9	US-09-899-303-12	Sequence 12, Appl	467	5	3.7	199	10	US-09-815-837-56	Sequence 56, Appl
395	5	3.7	158	9	US-09-899-303-12	Sequence 21, Appl	468	5	3.7	200	9	US-09-815-837-56	Sequence 26, Appl
396	5	3.7	158	9	US-09-899-303-12	Sequence 984, Ap	469	5	3.7	200	9	US-09-815-837-56	Sequence 26, Appl
397	5	3.7	159	9	US-09-973-025-10	Sequence 10, Appl	470	5	3.7	200	9	US-09-973-025-26	Sequence 160, Ap
398	5	3.7	159	9	US-09-973-025-10	Sequence 10, Appl	471	5	3.7	200	10	US-09-804-615-2	Sequence 2, Appl1
399	5	3.7	159	9	US-09-899-303-10	Sequence 6679, Ap	472	5	3.7	201	10	US-09-811-284-160	Sequence 160, Ap
400	5	3.7	159	10	US-09-220-920-12	Sequence 12, Appl	473	5	3.7	201	9	US-09-738-626-3806	Sequence 3806, Ap
401	5	3.7	159	10	US-09-220-920-12	Sequence 89, Appl	474	5	3.7	201	10	US-09-738-626-3806	Sequence 827, Ap
402	5	3.7	162	9	US-09-738-626-3598	Sequence 3598, Ap	475	5	3.7	202	10	US-09-764-868-827	Sequence 1156, Ap
403	5	3.7	162	9	US-09-738-626-3598	Sequence 429, Ap	476	5	3.7	203	9	US-09-925-301-1156	Sequence 206, Ap
404	5	3.7	162	9	US-09-738-626-4796	Sequence 429, Ap	477	5	3.7	204	10	US-09-712-364-206	Sequence 75, Appl
405	5	3.7	162	10	US-10-091-572-429	Sequence 800, Ap	478	5	3.7	205	10	US-09-828-644-75	Sequence 5128, Ap
406	5	3.7	166	9	US-09-867-550-800	Sequence 78, Appl	479	5	3.7	206	10	US-09-738-626-5128	Sequence 294, Ap
407	5	3.7	166	10	US-09-954-692-78	Sequence 78, Appl	480	5	3.7	206	10	US-09-912-020-294	Sequence 4727, Ap
408	5	3.7	166	10	US-09-559-671A-78	Sequence 89, Appl	481	5	3.7	207	10	US-09-887-855-5	Sequence 5, Appl1
409	5	3.7	166	10	US-09-815-242-5162	Sequence 5162, Ap	482	5	3.7	207	10	US-09-738-626-4727	Sequence 294, Ap
410	5	3.7	166	10	US-09-815-242-5162	Sequence 11642, A	483	5	3.7	207	10	US-09-881-752A-46	Sequence 4727, Ap
411	5	3.7	166	10	US-09-977-034-26	Sequence 26, Appl	484	5	3.7	208	9	US-09-764-864-980	Sequence 46, Appl
412	5	3.7	166	10	US-09-977-034-26	Sequence 28, Appl	485	5	3.7	208	9	US-09-989-293A-416	Sequence 980, Ap
413	5	3.7	167	10	US-10-239-028-2	Sequence 2, Appl1	486	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
414	5	3.7	167	10	US-09-977-034-27	Sequence 27, Appl	487	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
415	5	3.7	169	9	US-09-925-301-1470	Sequence 1470, Ap	488	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
416	5	3.7	170	10	US-10-222-162-16	Sequence 16, Appl	489	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
417	5	3.7	171	10	US-09-977-034-16	Sequence 16, Appl	490	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
418	5	3.7	171	10	US-10-011-548-18	Sequence 18, Appl	491	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
419	5	3.7	172	9	US-09-764-877-1567	Sequence 1567, Ap	492	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
420	5	3.7	172	9	US-09-738-626-5653	Sequence 5653, Ap	493	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
421	5	3.7	173	10	US-09-825-561A-86	Sequence 86, Appl	494	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
422	5	3.7	174	9	US-09-738-626-3756	Sequence 3756, Ap	495	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
423	5	3.7	174	10	US-09-864-761-38950	Sequence 38950, A	496	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
424	5	3.7	174	10	US-09-839-185-12	Sequence 12, Appl	497	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
425	5	3.7	176	8	US-08-424-5508-54	Sequence 54, Appl	498	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
426	5	3.7	176	9	US-10-116-252-11	Sequence 11, Appl	499	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
427	5	3.7	176	9	US-09-925-299-1085	Sequence 1085, Ap	500	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
428	5	3.7	176	10	US-09-925-299-1085	Sequence 7, Appl1	501	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
429	5	3.7	177	9	US-10-003-671A-7	Sequence 95, Appl	502	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
430	5	3.7	177	10	US-09-730-617-95	Sequence 95, Appl	503	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
431	5	3.7	177	10	US-09-764-864-993	Sequence 993, Appl	504	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
432	5	3.7	178	9	US-09-738-626-5509	Sequence 5509, Ap	505	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
433	5	3.7	178	9	US-10-222-162-18	Sequence 18, Appl	506	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
434	5	3.7	179	9	US-10-222-162-18	Sequence 39246, A	507	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
435	5	3.7	180	10	US-09-864-761-39246	Sequence 1001, Ap	508	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
436	5	3.7	181	9	US-09-764-868-11001	Sequence 40, Appl	509	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
437	5	3.7	181	10	US-09-220-920-40	Sequence 35721, A	510	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
438	5	3.7	181	10	US-09-864-761-35721	Sequence 35721, A	511	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
439	5	3.7	182	9	US-10-061-119-3	Sequence 43173, A	512	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
440	5	3.7	182	10	US-09-864-761-43173	Sequence 12316, A	513	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
441	5	3.7	182	10	US-09-815-242-12316	Sequence 29, Appl	514	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
442	5	3.7	183	10	US-09-784-810A-29	Sequence 65, Appl	515	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
443	5	3.7	183	10	US-09-739-907-65	Sequence 5100, Ap	516	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
444	5	3.7	183	10	US-09-815-242-5100	Sequence 250, Ap	517	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
445	5	3.7	184	9	US-09-893-737-250	Sequence 6, Appl1	518	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
446	5	3.7	184	9	US-09-741-233A-6	Sequence 49, Appl	519	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
447	5	3.7	184	10	US-09-862-540-49	Sequence 24, Appl1	520	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
448	5	3.7	184	10	US-09-037-460-2	Sequence 18, Appl	521	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
449	5	3.7	185	10	US-09-853-6258-18	Sequence 24, Appl	522	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
450	5	3.7	185	9	US-09-973-025-24	Sequence 41, Appl	523	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
451	5	3.7	185	9	US-09-899-303-24	Sequence 16, Appl	524	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
452	5	3.7	186	10	US-09-220-920-41	Sequence 23, Appl	525	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
453	5	3.7	187	9	US-09-944-160-16	Sequence 1424, Ap	526	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
454	5	3.7	187	10	US-10-222-162-23	Sequence 1233, Ap	527	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
455	5	3.7	188	10	US-09-764-864-1424	Sequence 3578, Ap	528	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
456	5	3.7	188	10	US-09-764-864-1424	Sequence 15, Appl	529	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap
457	5	3.7	190	9	US-09-738-626-3578	Sequence 15, Appl	530	5	3.7	208	9	US-09-989-293A-416	Sequence 416, Ap

531	5	3.7	208	9	US-09-991-854-416	Sequence 416, App	604	5	3.7	216	10	US-09-850-373-5	Sequence 5, App11
532	5	3.7	208	9	US-09-997-349-416	Sequence 416, App	605	5	3.7	216	10	US-09-815-837-45	Sequence 45, App1
533	5	3.7	208	9	US-09-997-440-416	Sequence 416, App	606	5	3.7	216	10	US-09-815-837-46	Sequence 46, App1
534	5	3.7	208	9	US-09-997-628-416	Sequence 416, App	607	5	3.7	218	9	US-09-738-626-3636	Sequence 13636, App
535	5	3.7	208	9	US-09-997-683-416	Sequence 416, App	608	5	3.7	218	9	US-10-278-173-150	Sequence 150, App
536	5	3.7	208	9	US-10-143-032-510	Sequence 510, App	609	5	3.7	219	10	US-09-815-242-5033	Sequence 5033, App
537	5	3.7	208	9	US-09-993-469-416	Sequence 416, App	610	5	3.7	219	10	US-09-815-242-1168	Sequence 12168, A
538	5	3.7	208	9	US-09-993-469-416	Sequence 416, App	611	5	3.7	219	10	US-09-815-242-11043	Sequence 11043, A
539	5	3.7	208	9	US-09-997-542-416	Sequence 416, App	612	5	3.7	220	9	US-09-949-842-20	Sequence 20, App1
540	5	3.7	208	9	US-10-123-108-510	Sequence 510, App	613	5	3.7	220	9	US-10-001-054-56	Sequence 56, App1
541	5	3.7	208	9	US-10-123-236-510	Sequence 510, App	614	5	3.7	220	10	US-09-220-920-26	Sequence 26, App1
542	5	3.7	208	9	US-10-123-261-510	Sequence 510, App	615	5	3.7	220	10	US-09-804-615-9	Sequence 9, App11
543	5	3.7	208	9	US-10-140-921-510	Sequence 510, App	616	5	3.7	220	10	US-09-815-242-11614	Sequence 11614, A
544	5	3.7	208	9	US-10-140-928-510	Sequence 510, App	617	5	3.7	220	10	US-09-815-837-47	Sequence 47, App1
545	5	3.7	208	9	US-09-990-427-416	Sequence 416, App	618	5	3.7	220	10	US-09-815-837-48	Sequence 48, App1
546	5	3.7	208	9	US-09-990-439-416	Sequence 416, App	619	5	3.7	221	10	US-09-925-300-1659	Sequence 1659, App
547	5	3.7	208	9	US-10-121-045-510	Sequence 510, App	620	5	3.7	222	9	US-10-092-925-2	Sequence 2, App11
548	5	3.7	208	9	US-10-123-292-510	Sequence 510, App	621	5	3.7	222	9	US-10-227-884-48	Sequence 48, App1
549	5	3.7	208	9	US-10-123-903-510	Sequence 510, App	622	5	3.7	222	9	US-10-230-163-48	Sequence 48, App1
550	5	3.7	208	9	US-10-124-819-510	Sequence 510, App	623	5	3.7	222	9	US-10-218-631-48	Sequence 48, App1
551	5	3.7	208	9	US-10-124-822-510	Sequence 510, App	624	5	3.7	222	9	US-10-230-338-48	Sequence 48, App1
552	5	3.7	208	9	US-10-140-925-510	Sequence 510, App	625	5	3.7	222	9	US-10-230-414-48	Sequence 48, App1
553	5	3.7	208	9	US-10-160-498-510	Sequence 510, App	626	5	3.7	222	9	US-10-216-1594-48	Sequence 48, App1
554	5	3.7	208	10	US-09-815-242-5302	Sequence 5302, App	627	5	3.7	222	9	US-10-218-849-48	Sequence 48, App1
555	5	3.7	208	10	US-09-989-722-416	Sequence 416, App	628	5	3.7	222	9	US-10-227-883-48	Sequence 48, App1
556	5	3.7	208	10	US-09-989-723-416	Sequence 416, App	629	5	3.7	222	9	US-10-227-883-48	Sequence 48, App1
557	5	3.7	208	10	US-09-989-279-416	Sequence 416, App	630	5	3.7	222	10	US-09-897-772-2	Sequence 2, App11
558	5	3.7	208	10	US-09-989-727-416	Sequence 416, App	631	5	3.7	222	10	US-09-815-242-5034	Sequence 5034, App
559	5	3.7	208	10	US-09-815-837-102	Sequence 102, App	632	5	3.7	223	9	US-09-811-088-8	Sequence 8, App11
560	5	3.7	208	10	US-09-989-731-416	Sequence 416, App	633	5	3.7	223	9	US-10-097-065-325	Sequence 325, App
561	5	3.7	208	10	US-09-989-732-416	Sequence 416, App	634	5	3.7	223	9	US-10-274-974-2	Sequence 2, App11
562	5	3.7	208	10	US-09-991-073-416	Sequence 416, App	635	5	3.7	223	10	US-09-816-494-5	Sequence 5, App11
563	5	3.7	208	10	US-09-990-442-416	Sequence 416, App	636	5	3.7	224	9	US-09-712-363-189	Sequence 189, App
564	5	3.7	208	10	US-09-991-163-416	Sequence 416, App	637	5	3.7	224	10	US-09-220-920-29	Sequence 29, App1
565	5	3.7	208	10	US-09-993-604-416	Sequence 416, App	638	5	3.7	224	10	US-09-804-615-16	Sequence 16, App1
566	5	3.7	208	10	US-09-990-436-416	Sequence 416, App	639	5	3.7	224	10	US-09-804-615-34	Sequence 34, App1
567	5	3.7	208	10	US-09-989-721-416	Sequence 416, App	640	5	3.7	225	9	US-09-738-626-3362	Sequence 3622, App
568	5	3.7	209	9	US-09-973-025-8	Sequence 8, App11	641	5	3.7	225	10	US-09-881-7522A-50	Sequence 50, App11
569	5	3.7	209	9	US-09-899-303-8	Sequence 8, App11	642	5	3.7	226	10	US-09-796-149-5	Sequence 5, App11
570	5	3.7	209	9	US-10-222-162-45	Sequence 45, App1	643	5	3.7	228	9	US-09-862-520-56	Sequence 56, App1
571	5	3.7	209	10	US-09-815-837-58	Sequence 58, App1	644	5	3.7	229	9	US-09-738-626-4779	Sequence 4779, App
572	5	3.7	210	9	US-09-973-025-28	Sequence 28, App1	645	5	3.7	230	9	US-09-738-626-4634	Sequence 4634, App
573	5	3.7	210	9	US-09-899-303-28	Sequence 28, App1	646	5	3.7	230	9	US-10-000-2556A-143	Sequence 143, App
574	5	3.7	210	10	US-09-815-837-40	Sequence 40, App1	647	5	3.7	232	9	US-09-738-626-4081	Sequence 4081, App
575	5	3.7	211	9	US-09-989-920-194	Sequence 194, App	648	5	3.7	232	10	US-09-925-302-755	Sequence 755, App
576	5	3.7	211	9	US-09-738-626-3723	Sequence 3723, App	649	5	3.7	232	10	US-09-738-626-6370	Sequence 6370, App
577	5	3.7	211	10	US-09-729-674-12	Sequence 12, App1	650	5	3.7	235	9	US-09-738-626-6379	Sequence 6379, App
578	5	3.7	211	10	US-09-765-205-30	Sequence 30, App1	651	5	3.7	235	9	US-09-533-029-74	Sequence 74, App1
579	5	3.7	211	10	US-09-815-837-39	Sequence 39, App1	652	5	3.7	235	10	US-09-938-803-5	Sequence 5, App11
580	5	3.7	211	10	US-09-815-837-59	Sequence 59, App1	653	5	3.7	235	10	US-09-764-864-1415	Sequence 1415, App
581	5	3.7	212	9	US-09-973-025-4	Sequence 4, App11	654	5	3.7	237	9	US-09-817-315-2	Sequence 2, App11
582	5	3.7	212	9	US-09-899-303-4	Sequence 4, App11	655	5	3.7	237	10	US-09-220-920-32	Sequence 32, App1
583	5	3.7	212	9	US-10-222-162-31	Sequence 31, App1	656	5	3.7	237	10	US-09-881-752A-56	Sequence 56, App1
584	5	3.7	212	9	US-10-222-162-32	Sequence 32, App1	657	5	3.7	238	9	US-09-813-398-40	Sequence 40, App1
585	5	3.7	212	9	US-10-222-162-43	Sequence 43, App1	658	5	3.7	239	9	US-09-973-025-22	Sequence 22, App1
586	5	3.7	212	9	US-10-222-162-44	Sequence 44, App1	659	5	3.7	239	9	US-09-899-303-22	Sequence 22, App1
587	5	3.7	212	10	US-09-815-837-37	Sequence 37, App1	660	5	3.7	239	9	US-09-925-299-834	Sequence 834, App
588	5	3.7	212	10	US-09-815-837-38	Sequence 38, App1	661	5	3.7	239	9	US-10-261-494-4	Sequence 4, App11
589	5	3.7	213	9	US-10-222-162-39	Sequence 39, App1	662	5	3.7	239	10	US-09-925-299-834	Sequence 834, App
590	5	3.7	213	10	US-09-815-837-41	Sequence 41, App1	663	5	3.7	239	10	US-09-840-787-102	Sequence 102, App1
591	5	3.7	213	10	US-09-815-837-43	Sequence 43, App1	664	5	3.7	239	10	US-09-800-728-102	Sequence 102, App1
592	5	3.7	213	10	US-09-815-837-44	Sequence 44, App1	665	5	3.7	239	10	US-09-925-300-1237	Sequence 1237, App
593	5	3.7	213	10	US-09-815-837-57	Sequence 57, App1	666	5	3.7	240	9	US-10-114-893-102	Sequence 102, App
594	5	3.7	213	10	US-09-765-298A-16	Sequence 16, App1	667	5	3.7	240	10	US-09-815-242-11656	Sequence 11656, A
595	5	3.7	214	9	US-10-222-162-35	Sequence 35, App1	668	5	3.7	241	9	US-10-104-580-12	Sequence 12, App1
596	5	3.7	214	10	US-09-867-550-944	Sequence 944, App	669	5	3.7	241	10	US-09-764-864-935	Sequence 935, App
597	5	3.7	214	10	US-09-815-837-42	Sequence 42, App1	670	5	3.7	243	10	US-09-864-761-49091	Sequence 49091, A
598	5	3.7	215	9	US-09-738-626-5708	Sequence 5708, App	671	5	3.7	243	10	US-09-925-301-1033	Sequence 1033, App
599	5	3.7	215	10	US-09-815-242-12955	Sequence 12955, A	672	5	3.7	244	9	US-10-260-877-84	Sequence 84, App1
600	5	3.7	215	10	US-09-815-242-13084	Sequence 13084, A	673	5	3.7	244	9	US-09-893-519A-17	Sequence 17, App1
601	5	3.7	215	10	US-09-938-719-6	Sequence 6, App11	674	5	3.7	244	10	US-09-925-301-1096	Sequence 1096, App
602	5	3.7	215	10	US-09-938-726-6	Sequence 6, App11	675	5	3.7	245	10	US-09-815-242-13255	Sequence 13255, A
603	5	3.7	215	10	US-09-938-703-6	Sequence 6, App11	676	5	3.7	246	9	US-09-738-626-5396	Sequence 5396, App

677	5	3.7	246	10	US-09-912-020-280	Sequence 280, App	750	5	3.7	282	9	US-09-905-291A-127	Sequence 127, App
678	5	3.7	249	9	US-09-989-442-147	Sequence 147, App	751	5	3.7	282	9	US-09-971-536-73	Sequence 73, App1
679	5	3.7	249	9	US-10-073-865-123	Sequence 123, App	752	5	3.7	282	9	US-10-051-325-9	Sequence 9, App1
680	5	3.7	249	9	US-10-083-357-754	Sequence 754, App	753	5	3.7	282	9	US-09-902-885-127	Sequence 127, App
681	5	3.7	249	10	US-09-764-853-771	Sequence 771, App	754	5	3.7	282	9	US-09-907-824-117	Sequence 127, App
682	5	3.7	250	9	US-09-738-626-6273	Sequence 6273, Ap	755	5	3.7	282	9	US-09-907-841-127	Sequence 127, App
683	5	3.7	251	9	US-09-738-626-5514	Sequence 5514, Ap	756	5	3.7	282	9	US-09-904-011-117	Sequence 127, App
684	5	3.7	251	10	US-09-925-302-504	Sequence 504, App	757	5	3.7	282	9	US-10-028-072-312	Sequence 312, App
685	5	3.7	252	10	US-09-350-874-20	Sequence 20, App1	758	5	3.7	282	9	US-10-028-072-312	Sequence 312, App
686	5	3.7	252	10	US-09-350-874-20	Sequence 22, App1	759	5	3.7	282	9	US-09-906-742-117	Sequence 117, App
687	5	3.7	252	10	US-09-350-874-28	Sequence 28, App1	760	5	3.7	282	9	US-10-121-049-312	Sequence 312, App
688	5	3.7	252	10	US-09-350-874-42	Sequence 42, App1	761	5	3.7	282	9	US-10-123-904-312	Sequence 312, App
689	5	3.7	253	9	US-09-880-748-1449	Sequence 1449, Ap	762	5	3.7	282	9	US-09-906-838-127	Sequence 127, App
690	5	3.7	254	10	US-09-942-858-6	Sequence 6, App1	763	5	3.7	282	9	US-09-907-613-127	Sequence 127, App
691	5	3.7	255	9	US-09-866-050A-706	Sequence 706, App	764	5	3.7	282	9	US-09-907-942-127	Sequence 127, App
692	5	3.7	257	1	US-08-882-431-2	Sequence 2, App1	765	5	3.7	282	9	US-10-140-470-312	Sequence 312, App
693	5	3.7	257	9	US-10-002-784A-2	Sequence 2, App1	766	5	3.7	282	9	US-10-176-921-312	Sequence 312, App
694	5	3.7	257	10	US-09-350-874-16	Sequence 16, App1	767	5	3.7	282	9	US-10-176-918-312	Sequence 312, App
695	5	3.7	260	9	US-09-738-626-3623	Sequence 3623, Ap	768	5	3.7	282	9	US-10-140-474-312	Sequence 312, App
696	5	3.7	261	9	US-10-043-649-2	Sequence 2, App1	769	5	3.7	282	9	US-09-904-820-127	Sequence 127, App
697	5	3.7	261	10	US-09-841-132-517	Sequence 517, App	770	5	3.7	282	9	US-09-904-859-127	Sequence 127, App
698	5	3.7	261	10	US-09-927-616A-6	Sequence 6, App1	771	5	3.7	282	9	US-09-902-402-312	Sequence 312, App
699	5	3.7	262	9	US-10-114-893-318	Sequence 318, App	772	5	3.7	282	9	US-09-902-903-127	Sequence 127, App
700	5	3.7	262	9	US-10-114-893-318	Sequence 8, App1	773	5	3.7	282	9	US-10-143-114-312	Sequence 312, App
701	5	3.7	262	10	US-09-815-242-11453	Sequence 11453, A	774	5	3.7	282	9	US-09-904-786-127	Sequence 127, App
702	5	3.7	262	10	US-09-223-047-8	Sequence 2, App1	775	5	3.7	282	9	US-09-906-646-127	Sequence 127, App
703	5	3.7	263	9	US-09-973-025-6	Sequence 6, App1	776	5	3.7	282	9	US-09-906-700-127	Sequence 127, App
704	5	3.7	263	9	US-09-893-303-6	Sequence 6, App1	777	5	3.7	282	9	US-10-140-002-312	Sequence 312, App
705	5	3.7	263	10	US-09-864-761-33879	Sequence 33879, A	778	5	3.7	282	9	US-09-902-903-127	Sequence 127, App
706	5	3.7	264	9	US-09-956-508A-3	Sequence 3, App1	779	5	3.7	282	9	US-09-903-749A-127	Sequence 127, App
707	5	3.7	264	9	US-10-215-759-18	Sequence 4, App1	780	5	3.7	282	9	US-09-903-786-127	Sequence 127, App
708	5	3.7	264	9	US-10-215-759-19	Sequence 19, App1	781	5	3.7	282	9	US-10-101-66A-771	Sequence 771, App
709	5	3.7	264	10	US-09-865-578-14	Sequence 14, App1	782	5	3.7	282	9	US-10-142-119-312	Sequence 312, App
710	5	3.7	264	10	US-09-772-114-9	Sequence 9, App1	783	5	3.7	282	9	US-09-902-736-127	Sequence 127, App
711	5	3.7	264	10	US-09-986-944-1	Sequence 1, App1	784	5	3.7	282	9	US-09-904-119-127	Sequence 127, App
712	5	3.7	265	10	US-09-772-114-8	Sequence 6, App1	785	5	3.7	282	9	US-09-907-794-127	Sequence 127, App
713	5	3.7	265	10	US-09-772-114-6	Sequence 8, App1	786	5	3.7	282	9	US-10-123-262-312	Sequence 312, App
714	5	3.7	266	9	US-10-043-487-245	Sequence 245, App	787	5	3.7	282	9	US-10-142-423-312	Sequence 312, App
715	5	3.7	266	10	US-09-393-634-15	Sequence 15, App1	788	5	3.7	282	9	US-09-902-692-127	Sequence 127, App
716	5	3.7	266	10	US-09-734-556-162	Sequence 162, App	789	5	3.7	282	9	US-09-903-520-127	Sequence 127, App
717	5	3.7	266	10	US-09-791-171-153	Sequence 153, App	790	5	3.7	282	9	US-09-904-462-127	Sequence 127, App
718	5	3.7	267	9	US-09-773-748-1	Sequence 1, App1	791	5	3.7	282	9	US-09-905-056-127	Sequence 127, App
719	5	3.7	268	9	US-09-738-626-4753	Sequence 4753, Ap	792	5	3.7	282	9	US-09-905-127	Sequence 127, App
720	5	3.7	269	10	US-09-801-196-2	Sequence 2, App1	793	5	3.7	282	9	US-09-904-462-127	Sequence 127, App
721	5	3.7	269	10	US-09-925-301-976	Sequence 976, App	794	5	3.7	282	9	US-10-141-755-312	Sequence 312, App
722	5	3.7	269	10	US-09-853-100A-2	Sequence 2, App1	795	5	3.7	282	9	US-10-121-050-312	Sequence 312, App
723	5	3.7	269	10	US-09-833-503A-2	Sequence 30, App1	796	5	3.7	282	9	US-10-121-050-312	Sequence 312, App
724	5	3.7	269	10	US-09-854-816-30	Sequence 14, App1	797	5	3.7	282	9	US-09-907-925-127	Sequence 127, App
725	5	3.7	270	10	US-09-350-874-14	Sequence 18, App1	798	5	3.7	282	9	US-09-904-462-127	Sequence 127, App
726	5	3.7	270	10	US-09-350-874-18	Sequence 6, App1	799	5	3.7	282	9	US-09-905-381-127	Sequence 127, App
727	5	3.7	271	9	US-09-974-298-6	Sequence 11, App1	800	5	3.7	282	9	US-09-909-064-127	Sequence 127, App
728	5	3.7	271	9	US-09-770-528-11	Sequence 11, App1	801	5	3.7	282	9	US-10-143-028-312	Sequence 312, App
729	5	3.7	271	9	US-09-738-626-3940	Sequence 3940, Ap	802	5	3.7	282	9	US-10-123-108-312	Sequence 312, App
730	5	3.7	271	9	US-10-011-931-1	Sequence 1, App1	803	5	3.7	282	9	US-10-123-236-312	Sequence 312, App
731	5	3.7	271	10	US-09-854-280-35	Sequence 25, App1	804	5	3.7	282	9	US-10-123-236-312	Sequence 312, App
732	5	3.7	271	10	US-09-815-242-10050	Sequence 10050, A	805	5	3.7	282	9	US-10-123-236-312	Sequence 312, App
733	5	3.7	271	10	US-09-840-707A-1	Sequence 1, App1	806	5	3.7	282	9	US-10-140-921-312	Sequence 312, App
734	5	3.7	271	10	US-09-854-208-25	Sequence 25, App1	807	5	3.7	282	9	US-10-140-928-312	Sequence 312, App
735	5	3.7	271	10	US-09-854-208-25	Sequence 7, App1	808	5	3.7	282	9	US-10-176-847-34	Sequence 34, App1
736	5	3.7	272	10	US-09-919-172-7	Sequence 6325, Ap	809	5	3.7	282	9	US-09-907-575-127	Sequence 127, App
737	5	3.7	272	10	US-09-925-297-661	Sequence 661, App	810	5	3.7	282	9	US-10-121-045-312	Sequence 312, App
738	5	3.7	273	9	US-09-738-626-6325	Sequence 6325, Ap	811	5	3.7	282	9	US-10-123-292-312	Sequence 312, App
739	5	3.7	274	9	US-09-993-525-2	Sequence 74, App1	812	5	3.7	282	9	US-10-123-292-312	Sequence 312, App
740	5	3.7	274	9	US-09-847-102A-74	Sequence 5792, App	813	5	3.7	282	9	US-10-124-819-312	Sequence 312, App
741	5	3.7	275	9	US-09-738-626-5792	Sequence 73, App1	814	5	3.7	282	9	US-10-124-819-312	Sequence 312, App
742	5	3.7	277	10	US-09-925-300-1491	Sequence 1491, Ap	815	5	3.7	282	9	US-10-140-925-312	Sequence 312, App
743	5	3.7	277	10	US-09-847-102A-73	Sequence 1, App1	816	5	3.7	282	9	US-10-140-925-312	Sequence 312, App
744	5	3.7	278	9	US-09-927-738-1	Sequence 189, App	817	5	3.7	282	10	US-09-909-320-127	Sequence 127, App
745	5	3.7	278	10	US-10-108-605-189	Sequence 559, App	818	5	3.7	282	10	US-09-909-088-127	Sequence 127, App
746	5	3.7	279	9	US-09-764-853-559	Sequence 159, App	819	5	3.7	282	9	US-10-043-487-313	Sequence 313, App
747	5	3.7	279	9	US-10-001-857-159	Sequence 71, App1	820	5	3.7	284	9	US-09-738-626-5434	Sequence 5434, Ap
748	5	3.7	280	9	US-09-800-321A-71	Sequence 56, App1	821	5	3.7	285	10	US-09-772-114-7	Sequence 7, App1
749	5	3.7	281	9	US-09-893-519A-56	Sequence 45, App1	822	5	3.7	285	10	US-09-756-983-24	Sequence 24, App1

823	5	3.7	286	9	US-10-016-668-5	Sequence 5, Appli	896	5	3.7	319	10	US-09-815-242-13897	Sequence 13897, A
824	5	3.7	285	9	US-10-231-013-9	Sequence 9, Appli	897	5	3.7	319	10	US-09-945-201-9	Sequence 9, Appli
825	5	3.7	287	10	US-09-815-242-13772	Sequence 13772, A	898	5	3.7	320	9	US-10-083-620A-17	Sequence 17, Appli
826	5	3.7	288	9	US-09-738-626-4503	Sequence 4503, A	899	5	3.7	321	9	US-10-012-507-1	Sequence 1, Appli
827	5	3.7	289	10	US-09-815-242-4948	Sequence 4948, Ap	900	5	3.7	321	10	US-09-925-297-530	Sequence 530, App
828	5	3.7	289	10	US-09-815-242-10097	Sequence 10097, A	901	5	3.7	322	9	US-09-995-225-18	Sequence 18, Appli
829	5	3.7	289	10	US-09-815-242-14056	Sequence 14056, A	902	5	3.7	322	9	US-09-995-225-10	Sequence 20, Appli
830	5	3.7	290	10	US-09-815-242-5743	Sequence 5743, Ap	903	5	3.7	322	9	US-09-738-626-6747	Sequence 6747, Ap
831	5	3.7	290	10	US-09-815-242-13182	Sequence 13182, A	904	5	3.7	323	9	US-09-792-793A-31	Sequence 31, Appli
832	5	3.7	290	10	US-09-903-814-16	Sequence 16, Appl	905	5	3.7	323	10	US-09-943-002-12	Sequence 12, Appli
833	5	3.7	291	9	US-09-736-457-333	Sequence 333, App	906	5	3.7	323	10	US-09-801-368-352	Sequence 352, App
834	5	3.7	291	9	US-09-902-941-333	Sequence 333, App	907	5	3.7	324	9	US-09-738-626-3593	Sequence 3593, App
835	5	3.7	291	9	US-09-849-626-333	Sequence 333, App	908	5	3.7	324	9	US-09-800-321A-72	Sequence 72, Appl
836	5	3.7	291	9	US-10-017-754-333	Sequence 333, App	909	5	3.7	324	9	US-09-810-321A-73	Sequence 73, Appli
837	5	3.7	291	10	US-09-865-578-13	Sequence 13, Appl	910	5	3.7	324	10	US-09-815-242-13422	Sequence 13422, A
838	5	3.7	291	10	US-09-853-625B-19	Sequence 19, Appl	911	5	3.7	325	9	US-09-738-626-3920	Sequence 3920, Ap
839	5	3.7	292	9	US-09-738-626-5836	Sequence 5836, Ap	912	5	3.7	325	10	US-09-815-242-3805	Sequence 3805, Ap
840	5	3.7	292	9	US-09-738-626-6173	Sequence 6173, Ap	913	5	3.7	325	10	US-09-815-242-12947	Sequence 12947, A
841	5	3.7	293	9	US-09-738-626-6009	Sequence 4009, Ap	914	5	3.7	326	9	US-09-991-211-1	Sequence 1, Appli
842	5	3.7	293	9	US-09-738-626-5474	Sequence 5474, Ap	915	5	3.7	326	9	US-09-337-946A-22	Sequence 22, Appl
843	5	3.7	293	10	US-09-815-242-10488	Sequence 10488, A	916	5	3.7	326	10	US-09-943-002-8	Sequence 8, Appli
844	5	3.7	293	10	US-09-815-242-11301	Sequence 11201, A	917	5	3.7	326	10	US-09-815-242-10573	Sequence 10573, A
845	5	3.7	293	10	US-09-815-242-12329	Sequence 12329, A	918	5	3.7	326	10	US-09-784-810A-11	Sequence 11, Appl
846	5	3.7	293	10	US-09-815-242-12742	Sequence 12742, A	919	5	3.7	327	10	US-09-945-201-6	Sequence 6, Appli
847	5	3.7	294	10	US-09-815-242-5225	Sequence 5225, Ap	920	5	3.7	327	10	US-09-925-300-1372	Sequence 1372, Ap
848	5	3.7	295	9	US-09-738-626-6160	Sequence 6160, App	921	5	3.7	328	9	US-10-164-433-4	Sequence 4, Appli
849	5	3.7	295	9	US-09-975-719-341	Sequence 341, App	922	5	3.7	329	8	US-08-945-038-4	Sequence 4, Appli
850	5	3.7	295	10	US-09-815-242-10666	Sequence 10666, A	923	5	3.7	329	10	US-09-815-242-13944	Sequence 13944, A
851	5	3.7	295	10	US-09-815-242-11836	Sequence 11836, A	924	5	3.7	330	9	US-09-800-321A-31	Sequence 31, Appli
852	5	3.7	296	9	US-09-789-054A-10	Sequence 10, Appli	925	5	3.7	330	10	US-09-826-508-12	Sequence 12, Appl
853	5	3.7	296	10	US-09-815-242-4883	Sequence 4883, Ap	926	5	3.7	330	10	US-09-815-242-13726	Sequence 13726, A
854	5	3.7	299	9	US-09-510-332-81	Sequence 81, Appl	927	5	3.7	330	10	US-09-750-373-24	Sequence 24, Appl
855	5	3.7	300	10	US-09-393-634-5	Sequence 5, Appli	928	5	3.7	330	10	US-09-920-068B-4	Sequence 4, Appli
856	5	3.7	301	9	US-09-925-300-1421	Sequence 1421, Ap	929	5	3.7	331	9	US-09-344-982-6	Sequence 6, Appli
857	5	3.7	301	9	US-09-738-626-6976	Sequence 6976, Ap	930	5	3.7	332	9	US-10-245-103-76	Sequence 76, Appli
858	5	3.7	301	9	US-09-746-660A-54	Sequence 54, Appl	931	5	3.7	332	9	US-10-245-107-76	Sequence 76, Appli
859	5	3.7	301	9	US-10-226-136-15	Sequence 15, Appl	932	5	3.7	332	9	US-10-245-143-76	Sequence 76, Appli
860	5	3.7	302	9	US-09-738-626-6862	Sequence 6862, Ap	933	5	3.7	332	9	US-10-245-771-76	Sequence 76, Appli
861	5	3.7	302	10	US-09-883-720-14	Sequence 14, Appl	934	5	3.7	332	9	US-10-245-851-76	Sequence 76, Appli
862	5	3.7	303	9	US-09-991-211-5	Sequence 5, Appli	935	5	3.7	332	9	US-10-245-883-76	Sequence 76, Appli
863	5	3.7	305	10	US-09-883-720-12	Sequence 12, Appli	936	5	3.7	332	9	US-10-237-535-76	Sequence 76, Appli
864	5	3.7	307	9	US-09-738-626-5911	Sequence 5911, Ap	937	5	3.7	332	9	US-10-238-183-76	Sequence 76, Appli
865	5	3.7	307	10	US-09-799-848-4	Sequence 4, Appli	938	5	3.7	332	9	US-10-238-283-76	Sequence 76, Appli
866	5	3.7	307	10	US-09-815-242-5155	Sequence 5155, Ap	939	5	3.7	332	9	US-10-238-370-76	Sequence 76, Appli
867	5	3.7	307	10	US-09-815-242-12309	Sequence 12309, A	940	5	3.7	332	9	US-10-245-055-76	Sequence 76, Appli
868	5	3.7	307	10	US-09-815-242-12916	Sequence 12916, A	941	5	3.7	332	9	US-10-245-147-76	Sequence 76, Appli
869	5	3.7	309	9	US-09-738-626-6039	Sequence 6039, Ap	942	5	3.7	332	9	US-10-245-730-76	Sequence 76, Appli
870	5	3.7	309	9	US-09-738-626-6953	Sequence 6953, Ap	943	5	3.7	332	9	US-10-245-739-76	Sequence 76, Appli
871	5	3.7	309	9	US-09-510-332-93	Sequence 93, Appl	944	5	3.7	332	9	US-10-246-210-76	Sequence 76, Appli
872	5	3.7	310	1	US-08-964-716-42	Sequence 42, Appl	945	5	3.7	333	9	US-09-738-626-5683	Sequence 5683, Ap
873	5	3.7	310	1	US-10-022-832-36	Sequence 36, Appl	946	5	3.7	333	10	US-09-815-242-5629	Sequence 5629, Ap
874	5	3.7	310	9	US-09-877-843-57	Sequence 57, Appl	947	5	3.7	333	10	US-09-815-242-10847	Sequence 10847, A
875	5	3.7	311	9	US-09-893-519A-54	Sequence 54, Appl	948	5	3.7	334	9	US-09-971-536-57	Sequence 57, Appl
876	5	3.7	312	9	US-09-789-054A-22	Sequence 22, Appl	949	5	3.7	334	9	US-09-789-054A-4	Sequence 4, Appli
877	5	3.7	312	9	US-09-510-332-109	Sequence 109, App	950	5	3.7	335	9	US-10-076-785-12	Sequence 12, Appl
878	5	3.7	312	10	US-09-815-242-5819	Sequence 5819, Ap	951	5	3.7	335	10	US-09-815-242-12251	Sequence 12251, A
879	5	3.7	312	10	US-09-801-368-438	Sequence 438, App	952	5	3.7	335	10	US-09-886-055-345	Sequence 345, App
880	5	3.7	313	9	US-10-012-507-2	Sequence 2, Appli	953	5	3.7	337	10	US-10-076-785-8	Sequence 8, Appli
881	5	3.7	313	10	US-09-864-761-35804	Sequence 35804, A	954	5	3.7	337	10	US-09-751-299-4	Sequence 4, Appli
882	5	3.7	314	9	US-09-738-626-6011	Sequence 6011, Ap	955	5	3.7	338	9	US-10-076-785-10	Sequence 10, Appl
883	5	3.7	314	10	US-09-815-242-10822	Sequence 10822, A	956	5	3.7	338	10	US-09-925-301-978	Sequence 978, App
884	5	3.7	315	10	US-09-942-858-12	Sequence 12, Appl	957	5	3.7	339	10	US-09-943-671-2	Sequence 2, Appli
885	5	3.7	315	10	US-09-864-761-36159	Sequence 36159, A	958	5	3.7	340	9	US-09-971-536-65	Sequence 65, Appl
886	5	3.7	315	9	US-09-898-416-8	Sequence 8, Appli	959	5	3.7	340	9	US-09-789-054A-6	Sequence 6, Appli
887	5	3.7	316	9	US-09-738-626-5667	Sequence 5667, Ap	960	5	3.7	340	9	US-10-033-297-79	Sequence 79, Appl
888	5	3.7	316	10	US-10-101-464A-941	Sequence 941, App	961	5	3.7	340	9	US-09-940-244-79	Sequence 79, Appl
889	5	3.7	316	10	US-09-826-508-24	Sequence 24, Appl	962	5	3.7	340	9	US-09-940-244-136	Sequence 136, App
890	5	3.7	317	9	US-09-815-242-11934	Sequence 11934, A	963	5	3.7	340	10	US-09-777-430A-27	Sequence 27, Appl
891	5	3.7	317	9	US-10-152-300-3	Sequence 3, Appli	964	5	3.7	340	10	US-09-816-248-15	Sequence 15, Appl
892	5	3.7	317	10	US-09-837-235-2	Sequence 36, Appli	965	5	3.7	341	9	US-09-738-626-3688	Sequence 3688, Ap
893	5	3.7	318	10	US-09-765-205-36	Sequence 36, Appl	966	5	3.7	342	9	US-09-964-008-1	Sequence 1, Appli
894	5	3.7	318	10	US-09-764-864-878	Sequence 878, App	967	5	3.7	342	9	US-09-964-008-3	Sequence 3, Appli
895	5	3.7	319	10	US-09-815-242-10048	Sequence 10048, A	968	5	3.7	342	9	US-09-964-008-3	Sequence 3, Appli

```

969 5 3 7 342 9 US-09-991-211-2 Sequence 2, Appll
970 5 3 7 342 9 US-09-991-211-10 Sequence 10, Appll
971 5 3 7 342 9 US-09-802-640-24 Sequence 24, Appll
972 5 3 7 342 10 US-09-835-922-2 Sequence 2, Appll
973 5 3 7 342 10 US-09-827-937A-2 Sequence 2, Appll
974 5 3 7 342 10 US-09-815-242-10430 Sequence 10,30, A
975 5 3 7 342 10 US-09-815-242-12104 Sequence 12104, A
976 5 3 7 342 10 US-09-815-242-14040 Sequence 14040, A
977 5 3 7 342 10 US-09-780-576-2 Sequence 2, Appll
978 5 3 7 343 10 US-09-939-521-8 Sequence 8, Appll
979 5 3 7 344 9 US-09-712-363-165 Sequence 165, Appll
980 5 3 7 344 10 US-09-925-297-461 Sequence 461, Appll
981 5 3 7 345 9 US-09-993-225-10 Sequence 10, Appll
982 5 3 7 346 10 US-09-751-299-2 Sequence 2, Appll
983 5 3 7 346 10 US-09-925-301-1071 Sequence 1071, Ap
984 5 3 7 346 10 US-09-815-242-13437 Sequence 13437, A
985 5 3 7 346 10 US-09-815-242-13656 Sequence 13656, A
986 5 3 7 347 9 US-09-866-050A-326 Sequence 326, App
987 5 3 7 347 10 US-09-925-301-1118 Sequence 1118, Ap
988 5 3 7 348 9 US-09-789-054A-20 Sequence 20, Appll
989 5 3 7 348 9 US-09-990-940-10 Sequence 10, Appll
990 5 3 7 348 10 US-09-884-430-2 Sequence 2, Appll
991 5 3 7 349 10 US-09-916-051-1 Sequence 1, Appll
992 5 3 7 351 9 US-10-116-255-20 Sequence 20, Appll
993 5 3 7 351 10 US-09-840-787-19 Sequence 19, Appll
994 5 3 7 351 10 US-09-888-358-4 Sequence 4, Appll
995 5 3 7 352 9 US-10-278-173-110 Sequence 110, App
996 5 3 7 352 9 US-10-087-714-2 Sequence 2, Appll
997 5 3 7 353 9 US-09-911-345-2 Sequence 2, Appll
998 5 3 7 353 10 US-09-815-242-5323 Sequence 5323, Ap
999 5 3 7 353 10 US-09-939-521-12 Sequence 12, Appll
1000 5 3 7 353 12 US-10-109-885-9 Sequence 9, Appll

```

ALIGNMENTS

```

RESULT 1
US-09-815-242-11068
Sequence 11068, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11068
LENGTH: 345

```

```

; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(345)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-11068

```

```

Query Match
Best Local Similarity 100.0%; Score 7; DB 10; Length 345;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 67 SESKAKI 73
DB 111 SESKAKI 117

```

```

RESULT 2
US-09-862-027-21
Sequence 21, Application US/09862027
Patent No. US20020142428A1
GENERAL INFORMATION:
APPLICANT: Hodge, Martin R.
TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof
FILE REFERENCE: 35800/234862
CURRENT APPLICATION NUMBER: US/09/862,027
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 09/345,473
PRIOR FILING DATE: 1999-06-30.
NUMBER OF SEQ ID NOS: 82
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21
LENGTH: 596
TYPE: PRT
ORGANISM: avian
US-09-862-027-21

```

```

Query Match
Best Local Similarity 100.0%; Score 7; DB 10; Length 596;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 33 ANLTSSL 39
DB 79 ANLTSSL 85

```

```

RESULT 3
US-10-097-065-562
Sequence 562, Application US/10097065
Publication No. US20030055236A1
GENERAL INFORMATION:
APPLICANT: Moore, Paul A. et al.
TITLE OF INVENTION: 110 Human Secreted Proteins
FILE REFERENCE: PZ021P1
CURRENT APPLICATION NUMBER: US/10/097,065
PRIOR FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: PCT/US98/27059
PRIOR FILING DATE: 1998-12-17
PRIOR APPLICATION NUMBER: 60/070,923
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,007
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,057
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,006
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,369
PRIOR FILING DATE: 1997-12-19
PRIOR APPLICATION NUMBER: 60/068,367
PRIOR FILING DATE: 1997-12-19
PRIOR APPLICATION NUMBER: 60/068,368
PRIOR FILING DATE: 1997-12-19
PRIOR APPLICATION NUMBER: 60/068,169

```

```

; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,053
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,064
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,054
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,008
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,365
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 562
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-065-562
```

```
Query Match          4.5%; Score 6; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 115 SLKPCM 120
    |||||
Db 11 SLKPCM 16
```

```

RESULT 4
US-09-864-761-48009
; Sequence 48009, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
```

```

; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48009
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL109657.7
; OTHER INFORMATION: EXPRESSED IN BR474, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.1
; OTHER INFORMATION: SWISSPROT HIT: P94461, EVALU 1.90e-01
; OTHER INFORMATION: EST_HUMAN HIT: BF509718.1, EVALU 7.00e-09
US-09-864-761-48009
```

```
Query Match          4.5%; Score 6; DB 10; Length 29;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 109 OKIVRH 114
    |||||
Db 5 OKIVRH 10
```

```

RESULT 5
US-09-864-761-35812
; Sequence 35812, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecm1ca-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
```



```
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO: 35812
;; LENGTH: 42
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC007250.2
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.97
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
;; OTHER INFORMATION: EXPRESSED IN BLOOD, SIGNAL = 1.3
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.95
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.95
;; OTHER INFORMATION: EST_HUMAN HIT: AA807691.1, EVALUATE 2.80e+00
US-09-864-761-35812
```

```
Query Match
Best Local Similarity 4.5%; Score 6; DB 10; Length 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 ANLTTS 38
Db 14 ANLTTS 19
```

```
RESULT 6
US-10-097-065-561
;; Sequence 561, Application US/10097065
;; Publication No. US20030055236A1
;; GENERAL INFORMATION:
;; APPLICANT: Moore, Paul A. et al.
;; FILE OF INVENTION: 110 Human Secreted Proteins
;; CURRENT APPLICATION NUMBER: US/10/097,065
;; CURRENT FILING DATE: 2002-03-14
;; PRIOR APPLICATION NUMBER: PCT/US98/27059
;; PRIOR FILING DATE: 1998-12-17
;; PRIOR APPLICATION NUMBER: 60/070,923
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,007
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,057
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,006
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,369
;; PRIOR FILING DATE: 1997-12-19
;; PRIOR APPLICATION NUMBER: 60/068,367
;; PRIOR FILING DATE: 1997-12-19
;; PRIOR APPLICATION NUMBER: 60/068,368
;; PRIOR FILING DATE: 1997-12-19
;; PRIOR APPLICATION NUMBER: 60/068,169
;; PRIOR FILING DATE: 1997-12-19
;; PRIOR APPLICATION NUMBER: 60/068,053
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,064
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,054
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,008
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,365
;; PRIOR FILING DATE: 1997-12-19
;; NUMBER OF SEQ ID NOS: 672
;; SOFTWARE: PatentIn Ver. 2.0
```

```
;; SEQ ID NO: 561
;; LENGTH: 70
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-097-065-561
```

```
Query Match
Best Local Similarity 4.5%; Score 6; DB 9; Length 70;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 115 SLKPCM 120
Db 20 SLKPCM 25
```

```
RESULT 7
US-09-815-242-4896
;; Sequence 4896, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE OF INVENTION: Prokaryotes
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 4896
;; LENGTH: 76
;; TYPE: PRT
;; ORGANISM: Enterococcus faecalis
US-09-815-242-4896
```

```
Query Match
Best Local Similarity 4.5%; Score 6; DB 10; Length 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 86 PEAVSL 91
Db 16 PEAVSL 21
```

```
RESULT 8
US-09-864-761-41016
;; Sequence 41016, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
```



```
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: Aecomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 41016
;; LENGTH: 77
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC015473.3
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.2
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.6
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.7
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.8
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.7
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.6
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.3
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4
;; OTHER INFORMATION: SWISSPROT HIT: P16522, EVALU8 8.30e+00
;; OTHER INFORMATION: EST_HUMAN HIT: H94215.1, EVALU8 2.00e-16
US-09-864-761-41016

Query Match 4.5%; Score 6; DB 10; Length 77;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;; FILE REFERENCE: Aecomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 42371
;; LENGTH: 77
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL157955.1
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.9
;; OTHER INFORMATION: EST_HUMAN HIT: H94215.1, EVALU8 2.00e-16
;; OTHER INFORMATION: SWISSPROT HIT: P16522, EVALU8 8.30e+00
US-09-864-761-42371

Query Match 4.5%; Score 6; DB 10; Length 77;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Db 53 LKPCML 58

RESULT 10

US-09-815-242-10863
; Sequence 10863, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes In
TITLE OF INVENTION: Prokaryotes

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 10863

LENGTH: 81

TYPE: PRT

ORGANISM: Enterococcus faecalis

US-09-815-242-10863

Query Match

Best Local Similarity 4.5%; Score 6; DB 10; Length 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 PEAVSL 91

Db 19 PEAVSL 24

RESULT 11

US-10-007-280A-238

; Sequence 238, Application US/10007280A

; Publication No. US20030059784A1

; GENERAL INFORMATION:

APPLICANT: Sun, Yongming

APPLICANT: Recipon, Herve

APPLICANT: Salceda, Susana

APPLICANT: Chenghua, Liu

TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro

FILE REFERENCE: DEX-0257

CURRENT APPLICATION NUMBER: US/10/007,280A

PRIOR FILING DATE: 2001-11-07

PRIOR APPLICATION NUMBER: US 60/246,640

PRIOR FILING DATE: 2000-11-08

NUMBER OF SEQ ID NOS: 238

SOFTWARE: PatentIn version 3.1

SEQ ID NO 238

LENGTH: 88

TYPE: PRT

ORGANISM: Homo sapien

US-10-007-280A-238

Query Match

Best Local Similarity 4.5%; Score 6; DB 9; Length 88;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 RSLSLKP 118

Db 65 RSLSLKP 70

RESULT 12

US-10-101-464A-499

; Sequence 499, Application US/10101464A

; Publication No. US20030046728A1

; GENERAL INFORMATION:

APPLICANT: Strabala, Timothy

APPLICANT: Nieuwenhuizen, Nicolaas

APPLICANT: Higgins, Colleen M.

TITLE OF INVENTION: Compositions Isolated from Plant Cells

TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling

FILE REFERENCE: 11000.1020C2

CURRENT APPLICATION NUMBER: US/10/101,464A

CURRENT FILING DATE: 2002-03-18

PRIOR APPLICATION NUMBER: 09/704,302

PRIOR FILING DATE: 2000-11-01

PRIOR APPLICATION NUMBER: 09/228,986

PRIOR FILING DATE: 1999-01-12

PRIOR APPLICATION NUMBER: 60/162,866

PRIOR FILING DATE: 1999-11-01

PRIOR APPLICATION NUMBER: PCT/US00/00724

PRIOR FILING DATE: 2000-01-11

NUMBER OF SEQ ID NOS: 989

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 499

LENGTH: 116

TYPE: PRT

ORGANISM: Eucalyptus grandis

US-10-101-464A-499

Query Match

Best Local Similarity 4.5%; Score 6; DB 9; Length 116;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 SLTELT 95

Db 24 SLTELT 29

RESULT 13

US-09-895-728-6

; Sequence 6, Application US/09895728

; Patent No. US20020123475A1

; GENERAL INFORMATION:

APPLICANT: Millennium Pharmaceuticals, Inc.

APPLICANT: Leiby, Kevin

APPLICANT: Spaltmann, Frank

APPLICANT: Cook, William

TITLE OF INVENTION: 32626, A NOVEL HUMAN

TITLE OF INVENTION: UDP-GLYCOSYLTRANSFERASE AND USES THEREOF

FILE REFERENCE: 38155-20018.00

CURRENT APPLICATION NUMBER: US/09/895,728

PRIOR FILING DATE: 2001-06-29

PRIOR APPLICATION NUMBER: US 60/215,749

PRIOR FILING DATE: 2000-06-30

NUMBER OF SEQ ID NOS: 6

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6

LENGTH: 131

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE: OTHER INFORMATION: Consensus amino acid

US-09-895-728-6

Query Match 4.5%; Score 6; DB 10; Length 131;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 MKLSE 68
|||||
Db 101 MKLSE 106

RESULT 14

US-10-101-464A-738
; Sequence 738, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions isolated from plant cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101.464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302.
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 738
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-738

Query Match 4.5%; Score 6; DB 9; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 NLRTSL 39
|||||
Db 13 NLRTSL 18

RESULT 15

US-09-764-864-1364
; Sequence 1364, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper.
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1364
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (136)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (142)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1364

Query Match 4.5%; Score 6; DB 10; Length 133;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 51 QLTGHQ 56
|||||
Db 95 QLTGHQ 100

RESULT 16

US-10-101-464A-89
; Sequence 89, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions isolated from plant cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101.464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-89

Query Match 4.5%; Score 6; DB 9; Length 154;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 90 SLTELT 95
|||||
Db 114 SLTELT 119

RESULT 17

US-09-854-133-125
; Sequence 125, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Monamath, Raedoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 125
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapien

US-09-854-133-125

Query Match
Best Local Similarity 4.5%; Score 6; DB 9; Length 195;
Matches 6; Conservative 100.0%; Pred. No. 2.5e+02;
Mismatches 0; Indels 0; Gaps 0;

OY 91 LTELTT 96
|||||
DB 83 LTELTT 88

RESULT 18

US-09-738-973-125
; Sequence 125, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Filing, Steven P.
; APPLICANT: Mohamach, Radooh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indrias, Carol Joseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliott, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738.973
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 125
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-738-973-125

Query Match
Best Local Similarity 4.5%; Score 6; DB 10; Length 195;
Matches 6; Conservative 100.0%; Pred. No. 2.5e+02;
Mismatches 0; Indels 0; Gaps 0;

OY 91 LTELTT 96
|||||
DB 83 LTELTT 88

RESULT 19
US-10-072-349-90
; Sequence 90, Application US/10072349
; Publication No. US20030054420A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P110C1
; CURRENT APPLICATION NUMBER: US/10/072.349
; CURRENT FILING DATE: 2002-02-11
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 90
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (2)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (8)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (132)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (149)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (166)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (184)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (189)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (208)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-072-349-90

Query Match
Best Local Similarity 4.5%; Score 6; DB 9; Length 210;
Matches 6; Conservative 100.0%; Pred. No. 2.7e+02;
Mismatches 0; Indels 0; Gaps 0;

OY 94 LTTAAI 99
|||||
DB 124 LTTAAI 129

RESULT 20
US-09-764-855-90
; Sequence 90, Application US/09764855
; Patent No. US20020119919A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P110
; CURRENT APPLICATION NUMBER: US/09/764.855
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (2)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (8)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (132)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (149)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (166)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (184)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE

LOCATION: (189)
OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (208)
OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
US-09-764-855-90

Query Match 4.5%; Score 6; DB 10; Length 210;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 LTTAAI 99
|||||
DB 124 LTTAAI 129

RESULT 21
US-09-738-626-5266

Sequence 5266, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENOH, AKIHITO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5266
LENGTH: 211
TYPE: PR1
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5266

Query Match 4.5%; Score 6; DB 9; Length 211;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 TTSLIK 41
|||||
DB 105 TTSLIK 110

RESULT 22
US-09-738-626-6335

Sequence 6335, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENOH, AKIHITO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6335
LENGTH: 261
TYPE: PR1
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6335

Query Match 4.5%; Score 6; DB 9; Length 261;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 NARTEV 107
|||||
DB 92 NARTEV 97

RESULT 23
US-09-911-346-2

Sequence 2, Application US/09911346
Patent No. US20020106323A1
GENERAL INFORMATION:
APPLICANT: Ni, Jian
Yu, Guo-Liang
Gentz, Reiner
Rosen, Craig A.
TITLE OF INVENTION: NATURAL KILLER CELL ENHANCED FACTOR-C
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
STUART & OLSTEIN
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: New Jersey
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/911,346
FILING DATE: 24-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/467,265
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-456
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-911-346-2

Query Match
Best Local Similarity 4.5%; Score 6; DB 10; Length 271;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 69 SKAKIS 74
DB 77 SKAKIS 82

RESULT 24
US-09-764-864-925
Sequence 925, Application US/09764864
Patent No. US20020132753A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PM223
CURRENT APPLICATION NUMBER: US/09/764,864
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: PatentIn ver. 2.0
SEQ ID NO 925
LENGTH: 276
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (198)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (203)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (229)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (237)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-925

Query Match
Best Local Similarity 4.5%; Score 6; DB 10; Length 276;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 QLTGHQ 56
DB 95 QLTGHQ 100

RESULT 25
US-09-738-626-4972
Sequence 4972, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: PatentIn ver. 2.0
SEQ ID NO 925

PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 4972
LENGTH: 287
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-4972

Query Match
Best Local Similarity 4.5%; Score 6; DB 9; Length 287;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 15 FTLASC 20
DB 5 FTLASC 10

RESULT 26
US-10-027-450-26
Sequence 26, Application US/10027450
Patent No. US20020102715A1
GENERAL INFORMATION:
APPLICANT: Falco, Saverio Carl
APPLICANT: Hitz, William D.
APPLICANT: Kinney, Anthony J.
APPLICANT: Cahoon, Rebecca E.
APPLICANT: Rafalski, J. Antoni
TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMES
FILE REFERENCE: BB-1126
CURRENT APPLICATION NUMBER: US/10/027,450
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 54
SOFTWARE: Microsoft Word Version 7.0A
SEQ ID NO 26
LENGTH: 297
TYPE: PRT
ORGANISM: Oryza sativa
US-10-027-450-26

Query Match
Best Local Similarity 4.5%; Score 6; DB 12; Length 297;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 69 SKAKIS 74
DB 57 SKAKIS 62

RESULT 27
US-09-510-332-9
Sequence 9, Application US/09510332
Publication No. US20030022278A1
GENERAL INFORMATION:
APPLICANT: Zuker, Charles S.
APPLICANT: Adler, Jon Elliot
APPLICANT: Ryba, Nick
APPLICANT: Mueller, Ken
APPLICANT: Hoon, Mark
TITLE OF INVENTION: The Regents of the University of California
FILE REFERENCE: 02307E-098010US
CURRENT APPLICATION NUMBER: US/09/510,332
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 172
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 9

```
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human T2R05 (hGR05)
US-09-510-332-9
```

```
Query Match          4.5%; Score 6; DB 9; Length 299;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 44 VKOTCQ 49
      |||||
Db 279 VKOTCQ 284
```

```
RESULT 28
US-09-393-634-43
; Sequence 43, Application US/09393634
; Patent No. US20020051997A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliott
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: SF, a No. US20020051997A1el Family of Taste Receptors
; FILE REFERENCE: 02307E-09800005
; CURRENT APPLICATION NUMBER: US/09/393,634
; CURRENT FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human GR05
US-09-393-634-43
```

```
Query Match          4.5%; Score 6; DB 10; Length 299;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 44 VKOTCQ 49
      |||||
Db 279 VKOTCQ 284
```

```
RESULT 29
US-10-027-450-15
; Sequence 15, Application US/10027450
; Patent No. US20020102715A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Hiltz, William D.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: PLANT BRANCHED CHAIN AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB-1126
; CURRENT APPLICATION NUMBER: US/10/027,450
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/063,423
; PRIOR FILING DATE: 1997 October 28
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 15
; LENGTH: 307
```

```
; TYPE: PRT
; ORGANISM: Oryza sativa
; US-10-027-450-15
```

```
Query Match          4.5%; Score 6; DB 12; Length 307;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 42 HAVKQT 47
      |||||
Db 72 HAVKQT 77
```

```
RESULT 30
US-09-801-368-286
; Sequence 286, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Call, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Saitama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.117
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 286
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-286
```

```
Query Match          4.5%; Score 6; DB 10; Length 309;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 29 KNSSAN 34
      |||||
Db 291 KNSSAN 296
```

```
RESULT 31
US-10-260-877-32
; Sequence 32, Application US/10260877
; Publication No. US20030021813A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Chovan, Linda E.
; APPLICANT: Hessler, Paul E.
; APPLICANT: Reich, Karl A.
; TITLE OF INVENTION: ESSENTIAL BACTERIA GENES AND GENOME
; TITLE OF INVENTION: 'SCANNING IN HAEMOPHILUS INFLUENZAE FOR THE IDENTIFICATION OF
; FILE REFERENCE: 6565 US. P1
; CURRENT APPLICATION NUMBER: US/10/260,877
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/09/649,145
; PRIOR FILING DATE: 2000-08-25
```

NUMBER OF SEQ ID NOS: 137
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 32
LENGTH: 311
TYPE: PRT
ORGANISM: H. Influenzae
US-10-260-877-32

Query Match
Best Local Similarity 4.5%; Score 6; DB 9; Length 311;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 90 SLELT 95
|||||
DB 174 SLELT 179

RESULT 32
US-09-964-899-7
Sequence 7, Application US/09964899
Patent No. US2002017446A1
GENERAL INFORMATION:
APPLICANT: Cohen, Dalia et al.
TITLE OF INVENTION: Identification of Genes Involved in
FILE REFERENCE: Alzheimer's Disease Using Drosophila Melanogaster
CURRENT APPLICATION NUMBER: US/09/964,899
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/236,893
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/298,309
PRIOR FILING DATE: 2001-06-14
NUMBER OF SEQ ID NOS: 53
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 332
TYPE: PRT
ORGANISM: Homo Sapien
US-09-964-899-7

Query Match
Best Local Similarity 4.5%; Score 6; DB 9; Length 332;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 71 AKISRT 76
|||||
DB 175 AKISRT 180

RESULT 33
US-09-815-242-5618
Sequence 5618, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5618
LENGTH: 336
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5618

Query Match
Best Local Similarity 4.5%; Score 6; DB 10; Length 336;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 90 SLELT 95
|||||
DB 241 SLELT 246

RESULT 34
US-09-815-242-12540
Sequence 12540, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12540
LENGTH: 336
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12540

Query Match
Best Local Similarity 4.5%; Score 6; DB 10; Length 336;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 90 SLELT 95
|||||
DB 241 SLELT 246

RESULT 35
US-10-116-016-37
; Sequence 37, Application US/10116016
; Publication No. US20030054379A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT208C1
; CURRENT APPLICATION NUMBER: US/10/116,016
; CURRENT FILING DATE: 2002-04-05
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-016-37

Query Match 4.5%; Score 6; DB 9; Length 343;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LSSESX 70
|||||
Db 294 LSSESX 299

RESULT 36
US-09-764-848-37
; Sequence 37, Application US/09764848
; Patent No. US2002007270A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT208
; CURRENT APPLICATION NUMBER: US/09/764,848
; CURRENT FILING DATE: 2001-01-17
; Prior Application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-848-37

Query Match 4.5%; Score 6; DB 10; Length 343;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LSSESX 70
|||||
Db 294 LSSESX 299

RESULT 37
US-09-104-063-2
; Sequence 2, Application US/09104063
; Patent No. US20020168356A1
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PF4A Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/104,063
FILING DATE: 24-June-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/701265
FILING DATE: 22-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/664228
FILING DATE: 06-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/076093
FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/810782
FILING DATE: 19-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: P0706P2C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-09-104-063-2

Query Match 4.5%; Score 6; DB 9; Length 350;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 118 PCMLET 123
|||||
Db 29 PCMLET 34

RESULT 38
US-09-782-980-83
; Sequence 83, Application US/09782980
; Patent No. US20020072089A1
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran M.
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Busfield, Samantha J.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Gu, Wei
; APPLICANT: White, David
; TITLE OF INVENTION: NOVEL ITALY, IOR-2, STRIFE, TRASH, BDFG, LRSG, AND
; TITLE OF INVENTION: STMT PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: INT-121CP
; CURRENT APPLICATION NUMBER: US/09/782,980
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: PCT/US00/02125
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 09/448,076
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: 09/276,400
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: 60/117,580
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: 09/014,195

;; PRIOR FILING DATE: 1998-01-27
;; PRIOR APPLICATION NUMBER: 09/014,348
;; PRIOR FILING DATE: 1998-01-27
;; PRIOR APPLICATION NUMBER: 09/086,892
;; PRIOR FILING DATE: 1998-05-29
;; PRIOR APPLICATION NUMBER: 09/296,208
;; PRIOR FILING DATE: 1999-04-21
;; PRIOR APPLICATION NUMBER: 09/063,950
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 09/561,381
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 09/561,810
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 09/087,121
;; PRIOR FILING DATE: 1998-05-29
;; PRIOR APPLICATION NUMBER: 09/672,721
;; PRIOR FILING DATE: 2000-09-28
;; PRIOR APPLICATION NUMBER: 09/049,799
;; NUMBER OF SEQ ID NOS: 176
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO: 83
;; LENGTH: 350
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-782-980-83

Query Match 4.5%; Score 6; DB 10; Length 350;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 118 PCMLET 123
DB 29 PCMLET 34

RESULT 39
US-09-884-430-7
;; Sequence 7, Application US/09884430
;; Patent No. US20020151046A1
;; GENERAL INFORMATION:
;; APPLICANT: Glucksmann, Maria Alexandra
;; APPLICANT: Santlago, Immaculada Siles
;; TITLE OF INVENTION: 52871, A NOVEL HUMAN G PROTEIN COUPLED
;; TITLE OF INVENTION: RECEPTOR AND USES THEREOF
;; FILE REFERENCE: MNI-165
;; CURRENT APPLICATION NUMBER: US/09/884,430
;; PRIOR FILING DATE: 2001-06-18
;; PRIOR APPLICATION NUMBER: USSN 60/212,331
;; PRIOR FILING DATE: 2000-06-16
;; PRIOR APPLICATION NUMBER: USSN 60/269,758
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 8
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 7
;; LENGTH: 350
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-884-430-7

Query Match 4.5%; Score 6; DB 10; Length 350;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 118 PCMLET 123
DB 29 PCMLET 34

RESULT 40
US-09-992-807-3
;; Sequence 3, Application US/09992807
;; Patent No. US20020127240A1

;; GENERAL INFORMATION:
;; APPLICANT: Haake, David A.
;; Shang, Ellen S.
;; TITLE OF INVENTION: Leptospiara MEMBRANE PROTEINS
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Robbins, Berliner & Carson
;; STREET: 201 N. Figueroa Street, 5th Floor
;; CITY: Los Angeles
;; STATE: California
;; COUNTRY: USA
;; ZIP: 90012-2628
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/992,807
;; FILING DATE: 16-Jan-2002
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/444,646
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Wong, Wean Khing
;; REGISTRATION NUMBER: 33,561
;; REFERENCE/DOCKET NUMBER: 5656-107
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (213) 977-1001
;; TELEFAX: (213) 977-1003
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 364 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-992-807-3

Query Match 4.5%; Score 6; DB 10; Length 364;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 92 TELTFA 97
DB 246 TELTFA 251

Search completed: April 28, 2003, 16:20:02
Job time : 56 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 16:11:33 ; Search time 17 Seconds
(without alignments)
757.765 Million cell updates/sec

Title: US-10-048-196-2

Perfect score: 134
Sequence: 1 MMKIIYVTAITMTAFITASC.....SLKPCMLETVNAFIVPTTR 134

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 1000 summaries

Database :

1: PIR_73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	6.0	100	2 AI3337	hypothetical prote
2	8	6.0	330	2 S37595	mucin UTL10 - huma
3	8	6.0	543	2 S35047	mucin JVL7 - huma
4	8	6.0	610	2 S35049	mucin JER57 - huma
5	8	6.0	3570	2 T45025	mucin MUC5B, trach
6	7	5.2	64	2 G97259	uncharacterized pr
7	7	5.2	132	2 AE3445	ATP synthase BME11
8	7	5.2	174	1 VCFYUR	coat protein gp37
9	7	5.2	174	2 AH3296	dUTP diphosphatase
10	7	5.2	175	2 B82380	purine-binding che
11	7	5.2	186	2 AH3404	hypothetical prote
12	7	5.2	198	2 C48613	env polyprotein TM
13	7	5.2	213	2 D96842	hypothetical prote
14	7	5.2	246	1 VCFY37	coat protein gp37
15	7	5.2	246	2 S23734	env polyprotein gp37
16	7	5.2	246	2 A60398	env polyprotein -
17	7	5.2	257	2 A63362	env polyprotein -
18	7	5.2	271	2 A71885	hypothetical prote
19	7	5.2	277	2 T31710	hypothetical prote
20	7	5.2	281	2 E64216	hypothetical prote
21	7	5.2	284	2 T09452	vdic protein VC214
22	7	5.2	302	2 D82112	tyra protein - Hel
23	7	5.2	320	2 T23674	hypothetical prote
24	7	5.2	323	2 A64054	selenophosphate sy
25	7	5.2	343	2 JC7183	cathepsin Q (EC 3.
26	7	5.2	345	2 C64082	ATP-binding protei
27	7	5.2	434	2 AB3099	hypothetical prote
28	7	5.2	436	2 G96187	thioblocin secret
29	7	5.2	459	2 T11084	NADH2 dehydrogenas

30	7	5.2	466	1 DCBYO	ornithine decarbox
31	7	5.2	486	2 D70154	replication initia
32	7	5.2	499	2 F69001	hypothetical prote
33	7	5.2	524	2 S35341	ketlin - fruit fly
34	7	5.2	594	2 H48613	env polyprotein pr
35	7	5.2	596	1 TVEVSA	env-sea polyprotei
36	7	5.2	601	2 I48613	env polyprotein pr
37	7	5.2	603	1 VCFEVR	env polyprotein -
38	7	5.2	630	2 S49659	probable membrane
39	7	5.2	692	2 T00025	PSD-95 binding pro
40	7	5.2	928	2 AC1312	ATP-dependent heli
41	7	5.2	977	2 T00014	DAP-1-alpha protei
42	7	5.2	1039	2 E90308	hypothetical prote
43	6	4.5	44	2 A53288	major pollen aller
44	6	4.5	57	2 D82927	hypothetical prote
45	6	4.5	65	2 S16463	hypothetical prote
46	6	4.5	65	2 S53704	hypothetical prote
47	6	4.5	86	2 E90905	hypothetical prote
48	6	4.5	89	2 S74731	ribosomal protein
49	6	4.5	93	2 T50425	hypothetical prote
50	6	4.5	96	2 AB0853	hypothetical prote
51	6	4.5	99	2 D81351	hypothetical prote
52	6	4.5	108	2 T25616	hypothetical prote
53	6	4.5	110	2 E71151	hypothetical prote
54	6	4.5	110	2 B70601	hypothetical prote
55	6	4.5	118	2 F90459	hypothetical prote
56	6	4.5	119	2 T22029	hypothetical prote
57	6	4.5	123	2 S00531	hypothetical prote
58	6	4.5	134	2 AB0784	insulin-like prote
59	6	4.5	136	2 S19479	conserved hypotet
60	6	4.5	137	2 S17195	hypothetical prote
61	6	4.5	138	2 A43957	insulin-like pepti
62	6	4.5	139	2 E96030	conserved hypotet
63	6	4.5	139	2 AG3453	hypothetical prote
64	6	4.5	142	2 A13492	enoyl-CoA hydratase
65	6	4.5	144	2 S01897	psib protein - Esc
66	6	4.5	145	2 T42187	psib protein - Esc
67	6	4.5	146	2 AE2154	hypothetical prote
68	6	4.5	146	2 S61393	hypothetical prote
69	6	4.5	146	2 F83814	Na+/H+ antiporter
70	6	4.5	147	2 T30606	hypothetical prote
71	6	4.5	148	2 F71730	hypothetical prote
72	6	4.5	148	2 B97732	hypothetical prote
73	6	4.5	151	2 S54048	ribosomal protein
74	6	4.5	153	2 B89882	conserved hypotet
75	6	4.5	159	2 D70302	hypothetical prote
76	6	4.5	161	2 D48909	G protein-coupled
77	6	4.5	161	2 G72711	hypothetical prote
78	6	4.5	162	2 C96635	hypothetical prote
79	6	4.5	163	2 AH1946	probable coatomer
80	6	4.5	171	2 G81304	hypothetical prote
81	6	4.5	172	2 B71464	probable lipoprote
82	6	4.5	172	2 C81726	hypothetical prote
83	6	4.5	179	2 F97324	conserved hypotet
84	6	4.5	182	2 A34647	uncharacterized co
85	6	4.5	185	2 E90395	pregnancy-specific
86	6	4.5	193	2 A53835	hypothetical prote
87	6	4.5	196	2 I50675	SEC63 protein comp
88	6	4.5	198	2 C87583	retinol-binding pr
89	6	4.5	199	1 VAHR	hypothetical prote
90	6	4.5	203	2 A59421	plasma retinol-bin
91	6	4.5	205	2 C96998	metalloproteinase
92	6	4.5	206	2 A75493	phosphatidylserine
93	6	4.5	208	2 C82622	hypothetical prote
94	6	4.5	211	2 S63664	anthranilate synth
95	6	4.5	214	2 A12371	hypothetical prote
96	6	4.5	222	2 G64931	hypothetical prote
97	6	4.5	222	2 A98933	ynic protein - Esc
98	6	4.5	222	2 AF0707	probable phosphata
99	6	4.5	222	2 E85781	probable hydrolase
100	6	4.5	222	2 A49599	polymerase-associ
101	6	4.5	222	2 B49599	polymerase-associ
102	6	4.5	222	2 S34563	gene M1 protein -

103	6	4.5	222	2	AE2003	hypothetical prote
104	6	4.5	223	2	AB2339	hypothetical prote
105	6	4.5	227	2	C84431	hypothetical prote
106	6	4.5	238	2	AG4156	Mg2+-transporting
107	6	4.5	239	2	S64327	probable membrane
108	6	4.5	240	2	AC3266	asparagine transpo
109	6	4.5	246	2	F63174	conserved hypothet
110	6	4.5	249	2	AH2207	hypothetical prote
111	6	4.5	252	2	F70711	probable membrane
112	6	4.5	252	2	T52160	hypothetical prote
113	6	4.5	254	2	C75612	phosphate ABC tran
114	6	4.5	254	2	S25281	glif protein precu
115	6	4.5	255	2	G82127	conserved hypothet
116	6	4.5	255	2	D86752	prophage p12 prote
117	6	4.5	256	2	B64631	conserved hypothet
118	6	4.5	258	2	G64992	hypothetical prote
119	6	4.5	258	2	D91017	hypothetical prote
120	6	4.5	258	2	S58661	hypothetical prote
121	6	4.5	258	2	S50921	ribosomal protein
122	6	4.5	260	2	E87392	conserved hypothet
123	6	4.5	261	2	E69455	conserved hypothet
124	6	4.5	261	2	E84267	glyoxalase fimport
125	6	4.5	261	2	T34909	hypothetical prote
126	6	4.5	261	2	AD1866	phospho-N-acetylmu
127	6	4.5	262	2	AB4967	hypothetical prote
128	6	4.5	267	2	AG0947	probable Deor-fam1
129	6	4.5	267	2	B87625	hypothetical prote
130	6	4.5	268	2	AF2470	potassium channel
131	6	4.5	269	1	A27067	calretinin - chick
132	6	4.5	269	2	G91239	probable DEOR-type
133	6	4.5	269	2	F86076	probable DEOR-type
134	6	4.5	269	2	S40828	probable transcrip
135	6	4.5	270	2	G69469	conserved hypothet
136	6	4.5	271	2	G01790	probable thiol-spe
137	6	4.5	272	2	S77125	phycocyanin alpha
138	6	4.5	273	2	C72328	transaminase B hom
139	6	4.5	273	2	B28928	pregnancy-specific
140	6	4.5	273	2	S32927	psab protein precu
141	6	4.5	273	2	AC0159	chaperone protein
142	6	4.5	273	2	B83318	hypothetical prote
143	6	4.5	273	2	H97584	glutamate racemase
144	6	4.5	273	2	AI2805	pregnancy-specific
145	6	4.5	275	2	A28928	ABC-type transport
146	6	4.5	280	2	T50583	2-dehydro-3-deoxyp
147	6	4.5	281	2	A83192	hypothetical prote
148	6	4.5	281	2	G86822	pregnancy-specific
149	6	4.5	282	2	C28928	hypothetical prote
150	6	4.5	282	2	T06630	protein T24H7.1 [1
151	6	4.5	286	2	H88175	conserved hypothet
152	6	4.5	286	2	A95875	hypothetical prote
153	6	4.5	287	2	AB6771	hypothetical prote
154	6	4.5	287	2	B75155	hypothetical prote
155	6	4.5	289	2	T47025	hypothetical prote
156	6	4.5	289	2	AH0235	hypothetical prote
157	6	4.5	291	2	G70605	probable dicarboxy
158	6	4.5	292	2	T35270	probable hydrolase
159	6	4.5	293	2	B91219	glucose-1-phosphat
160	6	4.5	293	2	D86065	probable AP2 domai
161	6	4.5	295	2	T00399	hypothetical prote
162	6	4.5	295	2	G87559	hypothetical prote
163	6	4.5	297	2	B96274	hypothetical prote
164	6	4.5	297	2	AD3010	hypothetical prote
165	6	4.5	299	2	A10681	probable regulator
166	6	4.5	300	1	RDB085	cytochrome-b5 redu
167	6	4.5	301	2	E97879	hypothetical prote
168	6	4.5	303	2	AE3073	hypothetical prote
169	6	4.5	304	2	AG1182	hypothetical prote
170	6	4.5	306	2	D75562	N-acetyl-gamma-glu
171	6	4.5	309	1	TVBARI	GTP-binding protei
172	6	4.5	310	2	E98113	probable transketol
173	6	4.5	311	2	F64046	delta(2)-isopenten
174	6	4.5	311	2	S43799	hypothetical prote
175	6	4.5	313	2	C81288	probable sugar-nuc
176	6	4.5	313	2	T15855	hypothetical prote
177	6	4.5	314	1	OXBP2L	hypothetical prote
178	6	4.5	315	2	S72472	type II site-speci
179	6	4.5	315	2	AH2259	type II site-speci
180	6	4.5	317	2	S40535	transaldolase (EC
181	6	4.5	317	2	H90629	transaldolase B [1
182	6	4.5	317	2	H85480	transaldolase B [1
183	6	4.5	317	2	AH0502	transaldolase B [1
184	6	4.5	317	2	AD0057	transaldolase (EC
185	6	4.5	319	2	T28888	hypothetical prote
186	6	4.5	320	2	S35007	modulation protein
187	6	4.5	320	2	A95947	phosphate uptake A
188	6	4.5	323	2	AD1963	hypothetical prote
189	6	4.5	324	2	G43354	hypothetical prote
190	6	4.5	326	2	F43354	pregnancy-specific
191	6	4.5	326	2	T52591	pregnancy-specific
192	6	4.5	329	2	E64698	ribose-phosphate d
193	6	4.5	329	2	G71821	polysialic acid ca
194	6	4.5	329	2	T05728	hypothetical prote
195	6	4.5	331	2	T13992	probable cadmium-t
196	6	4.5	331	2	UC4373	SF16 protein, pol1
197	6	4.5	332	2	B75286	glycerolaldehyde-3-p
198	6	4.5	332	2	B91039	hypothetical prote
199	6	4.5	332	2	D85883	probable cytochrom
200	6	4.5	332	2	A65015	probable cytochrom
201	6	4.5	333	2	A43354	yeah protein - Esc
202	6	4.5	333	2	AG2172	pregnancy-specific
203	6	4.5	335	2	H43354	hypothetical prote
204	6	4.5	336	2	G90421	pregnancy-specific
205	6	4.5	336	2	E89850	alcohol dehydrogen
206	6	4.5	336	2	H81785	glycerolaldehyde-3-p
207	6	4.5	336	2	B81210	conserved hypothet
208	6	4.5	336	2	S84717	hypothetical prote
209	6	4.5	337	2	C64233	glycerolaldehyde-3-p
210	6	4.5	337	2	S73737	hypothetical prote
211	6	4.5	337	2	AI2909	oligopeptide ABC t
212	6	4.5	338	2	H72428	GTP-binding protei
213	6	4.5	339	2	F97719	probable protein t
214	6	4.5	341	2	B91107	probable protein t
215	6	4.5	341	2	E85952	glycerolaldehyde-3-p
216	6	4.5	342	2	F70391	probable membrane
217	6	4.5	342	2	AD0275	sugar ABC transpor
218	6	4.5	343	2	D98213	NMDH2 dehydrogenas
219	6	4.5	344	2	T14018	H+-exporting ATPas
220	6	4.5	345	2	S55969	probable SF21 prot
221	6	4.5	347	2	A84579	interleukin-8 rece
222	6	4.5	350	2	A39445	probable iron tran
223	6	4.5	350	2	T04749	hypothetical prote
224	6	4.5	350	2	T34002	pregnancy-specific
225	6	4.5	352	2	I77374	iron(III)-transport
226	6	4.5	352	2	AB3650	cysteine proteinas
227	6	4.5	360	2	S57777	GTP-binding protei
228	6	4.5	360	2	D84952	polygalacturonase
229	6	4.5	363	2	S16998	3-dehydroquinolase
230	6	4.5	363	2	AG3002	hypothetical prote
231	6	4.5	364	2	S76068	oxidoreductase all
232	6	4.5	365	2	AF2442	hypothetical prote
233	6	4.5	365	2	F87616	CONSTANS homolog
234	6	4.5	366	2	T07835	hydrotgenase expres
235	6	4.5	367	2	E75031	probable ABC trans
236	6	4.5	367	2	C71357	CONSTANS homolog
237	6	4.5	368	2	T07836	aspartate kinase (
238	6	4.5	368	2	G75081	endo glucanase fam1
239	6	4.5	370	2	B97001	hypothetical prote
240	6	4.5	372	2	F86189	glycerolaldehyde-3-p
241	6	4.5	374	2	T07990	probable acyl-CoA
242	6	4.5	375	2	F83326	thiamin-binding pe
243	6	4.5	376	2	G97684	transposase - frui
244	6	4.5	376	2	S26856	UDP-GlcNAc 2-epime
245	6	4.5	376	2	D90004	hypothetical prote
246	6	4.5	377	2	T16147	3-dehydroquinolase
247	6	4.5	377	2	A99281	s
248	6	4.5	378	2	S72172	myoglobin - japane

249	6	4.5	378	2	A10152	probable bacteriop	322	6	4.5	459	2	T44201	hypothetical prote
250	6	4.5	379	2	S42529	Opaque-2-related p	323	6	4.5	459	2	T44014	tegument pp65/72k,
251	6	4.5	379	2	T46002	hypothetical prote	324	6	4.5	461	2	D72414	anthranilate synth
252	6	4.5	380	1	TVMSF	transforming prote	325	6	4.5	461	2	A97491	cysteinyl-tRNA syn
253	6	4.5	380	2	S71129	glyceralddehyde-3-p	326	6	4.5	461	2	AH2708	cysteinyl-tRNA syn
254	6	4.5	380	2	F98299	probable acyl-CoA	327	6	4.5	473	2	T45954	hypothetical prote
255	6	4.5	380	2	A12983	acyl-CoA dehydroge	328	6	4.5	475	2	S31927	dropy protein - hu
256	6	4.5	380	2	A13611	butyryl-CoA dehydr	329	6	4.5	475	2	D86450	hypothetical prote
257	6	4.5	381	1	TVWVJ	transforming prote	330	6	4.5	476	2	AE2126	pyruvate kinase [1
258	6	4.5	381	2	G64047	cystithionine gamm	331	6	4.5	479	2	F83291	probable outer mem
259	6	4.5	383	2	T00674	hypothetical prote	332	6	4.5	481	2	A38598	mannose-6-phosphat
260	6	4.5	384	2	S14450	probable transposa	333	6	4.5	481	2	B83201	N-acetylglucosamin
261	6	4.5	385	2	S17761	enoyl-l-acyl-carrie	334	6	4.5	484	2	JC7350	transferrin-bindin
262	6	4.5	386	2	B97411	33k chaperonin (he	335	6	4.5	488	2	F81003	archaeosine tRNA-r
263	6	4.5	387	2	A53586	albumin-binding pr	336	6	4.5	491	2	G84346	hypothetical prote
264	6	4.5	389	2	S36638	glycoprotein Epi-1	337	6	4.5	495	2	T04466	hypothetical prote
265	6	4.5	390	2	T49619	hypothetical prote	338	6	4.5	497	2	C82025	probable periplasm
266	6	4.5	391	2	A43299	N-acetylneuraminic	339	6	4.5	498	2	H85040	hypothetical prote
267	6	4.5	395	2	D43354	pregnancy-specific	340	6	4.5	499	2	AD1499	hypothetical secre
268	6	4.5	395	2	F71711	acyl-CoA desaturas	341	6	4.5	499	2	AF3026	hypothetical prote
269	6	4.5	397	2	E91296	probable thymidine	342	6	4.5	501	2	JH0447	alpha-1a-adrenergic
270	6	4.5	397	2	C43354	pregnancy-specific	343	6	4.5	502	2	F85715	hypothetical prote
271	6	4.5	400	2	D83892	transcription regu	344	6	4.5	502	2	H81273	probable amino aci
272	6	4.5	406	2	E43354	pregnancy-specific	345	6	4.5	505	2	F82075	conserved hypochet
273	6	4.5	409	2	F90825	probable integrase	346	6	4.5	505	2	C98258	ABC transporter pr
274	6	4.5	410	2	S18570	hypothetical prote	347	6	4.5	506	2	G86806	lipopolysaccharide
275	6	4.5	412	2	H81825	polygalacturonase	348	6	4.5	507	2	B69957	conserved hypochet
276	6	4.5	412	2	S43340	membrane fusion pr	349	6	4.5	512	2	JH0098	anthranilate synth
277	6	4.5	414	2	UC7270	glyceralddehyde-3-p	350	6	4.5	516	2	G84442	probable nucleosid
278	6	4.5	414	2	E75129	G-protein-coupled	351	6	4.5	516	2	D84087	epidermal surface
279	6	4.5	416	2	S45484	hypothetical prote	352	6	4.5	522	2	S62941	probable membrane
280	6	4.5	416	2	G81258	glyceralddehyde-3-p	353	6	4.5	526	1	TVFVMT	protein-tyrosine k
281	6	4.5	416	2	A85684	serine transporter	354	6	4.5	526	1	TVFV60	protein-tyrosine k
282	6	4.5	416	2	E82231	probable integrase	355	6	4.5	526	1	TVFV60	protein-tyrosine k
283	6	4.5	416	2	A28277	NiS-related prote	356	6	4.5	526	1	OKFVYR	protein-tyrosine k
284	6	4.5	417	2	S29506	pregnancy-specific	357	6	4.5	526	2	S15582	protein-tyrosine k
285	6	4.5	418	2	A33258	neurotensin recept	358	6	4.5	526	2	S20808	protein-tyrosine k
286	6	4.5	419	2	A31135	pregnancy-specific	359	6	4.5	526	2	S26420	protein-tyrosine k
287	6	4.5	419	2	S56073	pregnancy-specific	360	6	4.5	526	2	JC4101	aspartate ammonia-
288	6	4.5	421	2	D86240	opaque-2 protein -	361	6	4.5	527	2	G89999	conserved hypochet
289	6	4.5	421	2	A44502	hypothetical prote	362	6	4.5	529	1	VGNZSP	DNA-directed RNA p
290	6	4.5	422	1	T24561	nucleoprotein - Ch	363	6	4.5	532	2	S23207	cell fusion glycop
291	6	4.5	422	2	JH0164	neurotensin recept	364	6	4.5	532	1	B55066	tyrosine decarboxy
292	6	4.5	424	2	B33258	pregnancy-specific	365	6	4.5	532	1	B34104	protein-tyrosine k
293	6	4.5	426	2	F75394	pregnancy-specific	366	6	4.5	533	1	TVCHS	protein-tyrosine k
294	6	4.5	426	2	A35964	pregnancy-specific	367	6	4.5	534	2	A37483	protein-tyrosine k
295	6	4.5	426	2	A35341	pregnancy-specific	368	6	4.5	534	2	T04663	F protein - Murray
296	6	4.5	426	2	F75394	hypothetical prote	369	6	4.5	536	2	T33480	hypothetical prote
297	6	4.5	427	2	T48303	meiosis-specific-1	370	6	4.5	536	2	B84772	hypothetical prote
298	6	4.5	428	2	A27658	pregnancy-specific	371	6	4.5	537	2	T33480	probable DnaJ prot
299	6	4.5	428	2	J50032	hypothetical prote	372	6	4.5	537	2	VGNZMU	cell fusion glycop
300	6	4.5	428	2	S57486	probable transcrip	373	6	4.5	538	1	VGNZMM	cell fusion glycop
301	6	4.5	429	2	A75156	ornithine decarbox	374	6	4.5	538	1	VGNZMS	cell fusion glycop
302	6	4.5	430	2	AB0811	two-component hypr	375	6	4.5	538	1	B60004	cell fusion glycop
303	6	4.5	431	2	S64704	ornithine decarbox	376	6	4.5	538	2	S52472	cell fusion glycop
304	6	4.5	431	2	A19799	avermectin-sensiti	377	6	4.5	539	2	S56287	cell fusion glycop
305	6	4.5	433	2	T03035	nifS homolog - fls	378	6	4.5	541	1	A43610	J-domain protein D
306	6	4.5	433	2	S50865	acetyl-CoA C-acylt	379	6	4.5	542	1	TVHUSC	protein-tyrosine k
307	6	4.5	434	2	T37521	transcription regu	380	6	4.5	545	2	S52313	protein-tyrosine k
308	6	4.5	436	2	AB0335	thymidine phosphor	381	6	4.5	546	2	S52314	protein-tyrosine k
309	6	4.5	439	2	C70396	thymidine phosphor	382	6	4.5	547	2	H87679	sensor histidine k
310	6	4.5	440	1	S56606	triglycer factor [lm	383	6	4.5	548	2	T27542	hypothetical prote
311	6	4.5	440	2	G66137	hypothetical prote	384	6	4.5	551	1	VGNZP2	probable amino aci
312	6	4.5	445	2	T31581	sterile apetala [v	385	6	4.5	551	1	VGNZPG	cell fusion glycop
313	6	4.5	445	2	B97863	probable signal pe	386	6	4.5	551	2	AC2311	cell fusion glycop
314	6	4.5	445	2	T21744	protein containing	387	6	4.5	551	2	S65289	hypothetical prote
315	6	4.5	446	2	T51368	hypothetical prote	388	6	4.5	553	1	VGNZNV	hypothetical prote
316	6	4.5	446	2	T10024	anthranilate synth	389	6	4.5	553	1	VGNZGB	cell fusion glycop
317	6	4.5	449	2	H97249	proline/betaine tr	390	6	4.5	553	1	D46329	cell fusion glycop
318	6	4.5	454	2	C96648	UDP-N-acetylglucos	391	6	4.5	553	1	E46329	cell fusion glycop
319	6	4.5	454	2	A11278		392	6	4.5	553	1	G46329	cell fusion glycop
320	6	4.5	455	2	H97830		393	6	4.5	553	1		cell fusion glycop
321	6	4.5	457	2	C82720		394	6	4.5	553	1	VGNZTE	cell fusion glycop

395	6	4.5	553	1	A36830	cell fusion glycop	468	6	4.5	784	2	C88558	protein ZK1098.3 l
396	6	4.5	553	1	B36830	cell fusion glycop	469	6	4.5	795	2	T47964	hypothetical prote
397	6	4.5	553	1	VGNZND	cell fusion glycop	470	6	4.5	795	2	B97294	stage II sporulat
398	6	4.5	553	1	I46329	cell fusion glycop	471	6	4.5	807	2	T24110	hypothetical prote
399	6	4.5	553	1	VGNZU1	cell fusion glycop	472	6	4.5	815	2	T40524	hypothetical prote
400	6	4.5	553	1	A46325	cell fusion glycop	473	6	4.5	823	2	D86165	protein F15K3.3 (1
401	6	4.5	553	1	B46329	cell fusion glycop	474	6	4.5	830	2	T56940	factor arrest prot
402	6	4.5	553	1	H46329	cell fusion glycop	475	6	4.5	847	2	T12977	hypothetical prote
403	6	4.5	553	2	S06345	cell fusion glycop	476	6	4.5	857	1	QOBE1L	glycoprotein B - h
404	6	4.5	553	2	S23620	F protein - Newcas	477	6	4.5	859	2	AC2089	adenylate cyclase
405	6	4.5	553	2	S38784	gene F protein - N	478	6	4.5	880	2	F75103	conserved hypothet
406	6	4.5	553	2	S23621	gene F protein - N	479	6	4.5	886	2	A87093	valyl-tRNA synthas
407	6	4.5	553	2	S38786	gene F protein - N	480	6	4.5	901	1	A49856	valine-tRNA ligase
408	6	4.5	553	2	S38785	gene F protein - N	481	6	4.5	905	2	AC2680	hypothetical prote
409	6	4.5	553	2	S23622	gene F protein - N	482	6	4.5	916	2	A97462	hypothetical prote
410	6	4.5	553	2	S40163	cell fusion protei	483	6	4.5	923	2	E83574	hypothetical prote
411	6	4.5	557	1	TFVGS2	protein-tyrosine k	484	6	4.5	928	2	AC1684	hypothetical prote
412	6	4.5	557	1	D83478	gamma-glutamyltran	485	6	4.5	930	2	D71617	ATP-dependent hali
413	6	4.5	560	2	A38731	alpha-1A adrenerg	486	6	4.5	932	2	H86325	SENA antigen/papal
414	6	4.5	561	1	VGNZ41	cell fusion glycop	487	6	4.5	942	2	T13014	cytochrome b245 be
415	6	4.5	562	1	T52481	cytochrome-c oxida	488	6	4.5	949	2	E82068	valyl-tRNA synthet
416	6	4.5	568	1	TFVGS1	protein-tyrosine k	489	6	4.5	953	2	S54020	probable membrane
417	6	4.5	568	1	D90525	hypothetical prote	490	6	4.5	954	2	G64121	valine-tRNA ligase
418	6	4.5	570	2	S24459	hypothetical prote	491	6	4.5	956	2	B71250	valine-tRNA ligase
419	6	4.5	571	2	T10232	hypothetical prote	492	6	4.5	957	2	S44748	C06G4.1 protein -
420	6	4.5	572	2	I39369	alpha-1A-adrenerg	493	6	4.5	967	2	S66852	hypothetical prote
421	6	4.5	576	2	AC2195	hypothetical prote	494	6	4.5	975	2	I59422	rsce8 - rat (fragm
422	6	4.5	580	2	B38418	jockey protein 1 -	495	6	4.5	982	1	S45444	BEM1 protein-blnd1
423	6	4.5	581	2	S63183	CMN1 protein - yea	496	6	4.5	982	1	S00954	pol polyprotein -
424	6	4.5	581	2	T04864	probable serine/th	497	6	4.5	982	1	S05954	conserved large me
425	6	4.5	588	1	TFVPR	protein-tyrosine k	498	6	4.5	983	2	A87063	env polyprotein pr
426	6	4.5	588	2	T48766	probable sugar tra	499	6	4.5	990	1	G46335	matng-type switc
427	6	4.5	594	1	A56684	acetylhydroxy acid	500	6	4.5	993	2	S21964	myosin heavy chain
428	6	4.5	605	2	G72238	lipopolysaccharide	501	6	4.5	1002	1	JQ0151	probable transmem
429	6	4.5	606	2	S35427	env polyprotein -	502	6	4.5	1003	2	T16740	hypothetical prote
430	6	4.5	613	2	T28952	hypothetical prote	503	6	4.5	1008	2	S55603	DNA polymerase rep
431	6	4.5	614	2	T42649	hypothetical prote	504	6	4.5	1008	2	S72698	transport protein .
432	6	4.5	626	1	A48648	acetylacetate synth	505	6	4.5	1008	2	F82263	probable multidrug
433	6	4.5	630	2	T00351	hypothetical prote	506	6	4.5	1036	2	AG1326	alpha-mannosidase
434	6	4.5	638	2	S67605	hypothetical prote	507	6	4.5	1052	2	T50127	hypothetical prote
435	6	4.5	638	2	T38863	hypothetical prote	508	6	4.5	1078	2	T18352	protein P120 - Myc
436	6	4.5	642	1	SYRPAI	5-aminolevulinate	509	6	4.5	1109	2	T18536	receptor-like prot
437	6	4.5	645	2	C64879	molR_2 protein - E	510	6	4.5	1111	2	T23047	UL37 protein - hum
438	6	4.5	653	2	S67035	probable membrane	511	6	4.5	1112	1	WMBEH7	sensory box sensor
439	6	4.5	663	2	T26835	hypothetical prote	512	6	4.5	1131	2	T31911	hypothetical prote
440	6	4.5	667	1	VCLJGL	env polyprotein pr	513	6	4.5	1132	2	G83776	hypothetical prote
441	6	4.5	670	2	A85819	hypothetical prote	514	6	4.5	1137	2	D75429	hypothetical prote
442	6	4.5	673	2	AB1936	conserved hypotet	515	6	4.5	1162	2	T31911	cuti4 protein - fl
443	6	4.5	680	2	B83154	mitosis initiate	516	6	4.5	1172	2	S51623	hypothetical prote
444	6	4.5	680	2	A97331	membrane associa	517	6	4.5	1181	2	C86349	hypothetical prote
445	6	4.5	695	2	T13648	mitosis initiate	518	6	4.5	1196	2	H86389	hypothetical prote
446	6	4.5	706	2	T49700	related to Ap-1-11	519	6	4.5	1234	2	T30160	hypothetical prote
447	6	4.5	707	2	C90999	probable terminase	520	6	4.5	1237	2	AC1583	internalin prote
448	6	4.5	709	2	A38436	mitosis initiation	521	6	4.5	1240	2	T03097	CD0 protein - huma
449	6	4.5	709	2	S40926	hypothetical prote	522	6	4.5	1253	2	T45787	disease resistance
450	6	4.5	713	2	B86315	F2H15.20 protein -	523	6	4.5	1256	2	T03096	CD0 protein - rat
451	6	4.5	715	2	H84799	hypothetical prote	524	6	4.5	1266	2	S14555	botulinum neurotox
452	6	4.5	726	2	H86205	hypothetical prote	525	6	4.5	1276	2	S70582	hypothetical prote
453	6	4.5	727	2	S27043	neurotransmitter t	526	6	4.5	1285	2	T24480	hypothetical prote
454	6	4.5	727	2	I56506	Na+/Cl(-)-depend	527	6	4.5	1298	2	S07245	hypothetical prote
455	6	4.5	730	2	T13792	NADH2 dehydrogen	528	6	4.5	1335	2	T17508	hypothetical prote
456	6	4.5	730	2	S44617	C50C3.11 protein -	529	6	4.5	1337	2	T38949	hypothetical prote
457	6	4.5	734	2	S4617	hypothetical prote	530	6	4.5	1337	2	T38949	hypothetical prote
458	6	4.5	739	2	B86816	phosphatidylinform	531	6	4.5	1342	2	A31946	serine/threonine-s
459	6	4.5	739	2	T51701	antitransilact phosp	532	6	4.5	1388	2	S70633	DNA-directed RNA p
460	6	4.5	745	2	T48366	subtilisin-like pr	533	6	4.5	1400	2	AH3345	probable autotrans
461	6	4.5	755	2	T48353	chemotaxis protein	534	6	4.5	1430	2	AF0351	probable ATP-depen
462	6	4.5	762	2	H87302	protein transport	535	6	4.5	1448	2	F83237	B. subtilis yuka p
463	6	4.5	765	2	T40674	glucan 1,4-alpha-g	536	6	4.5	1498	2	AF1082	B. subtilis yuka p
464	6	4.5	774	2	A70010	NADH dehydrogenase	537	6	4.5	1498	2	AF1082	hypothetical prote
465	6	4.5	776	2	T52118	probable replicati	538	6	4.5	1506	2	E86302	Integumentary muc1
466	6	4.5	776	2	T52118	glucan 1,4-alpha-g	539	6	4.5	1506	2	T30886	
467	6	4.5	778	1	ALBYG		540	6	4.5	1506	2	T30886	

541	6	4.5	1516	2	T01055	hypothetical prote
542	6	4.5	1608	2	A28182	hemolysin A - Serr
543	6	4.5	1636	2	S60403	probable membrane
544	6	4.5	1661	2	S64800	IgA-specific metal
545	6	4.5	1668	2	A60272	hypothetical prote
546	6	4.5	1756	2	T07566	hypothetical prote
547	6	4.5	1787	2	A95273	conserved hypochet
548	6	4.5	1815	2	B95942	conserved hypochet
549	6	4.5	1818	2	AE3011	probable nuclear p
550	6	4.5	1837	2	T41023	myosin-like protei
551	6	4.5	1875	2	S38173	beta-galactosidase
552	6	4.5	1928	2	JS0610	gene 11-1 protein
553	6	4.5	1948	2	S00485	coagulation factor
554	6	4.5	2183	2	T42764	ALR protein homolo
555	6	4.5	2422	2	T12687	probable membrane
556	6	4.5	2493	2	S45734	hypothetical prote
557	6	4.5	2526	2	T20531	hypothetical prote
558	6	4.5	2722	2	T20532	protein F07A11.6 [
559	6	4.5	2738	2	E88320	adherence factor T
560	6	4.5	3225	2	D81702	apollipophorin prec
561	6	4.5	3305	2	T18358	breast cancer tumo
562	6	4.5	3418	1	G02334	367k tegument prot
563	6	4.5	3421	1	WZBBB6	genome polypeptid
564	6	4.5	3968	1	GNWVBV	hypothetical prote
565	6	4.5	4307	2	T20721	dynein heavy chain
566	6	4.5	4639	1	A54794	hypothetical prote
567	6	4.5	6713	2	B89921	type I fatty acid
568	6	4.5	8243	2	T31307	T-cell receptor be
569	5	3.7	16	2	PH0749	34k protein - rape
570	5	3.7	18	2	S28424	hypothetical prote
571	5	3.7	26	2	S12146	leu operon leader
572	5	3.7	28	1	LEBBLT	NADH2 dehydrogenas
573	5	3.7	28	2	AG0516	NADH2 dehydrogenas
574	5	3.7	29	2	T12242	NADH2 dehydrogenas
575	5	3.7	29	2	T12246	NADH2 dehydrogenas
576	5	3.7	29	2	T17079	NADH2 dehydrogenas
577	5	3.7	29	2	T17076	ribosomal protein
578	5	3.7	32	2	S51061	hypothetical prote
579	5	3.7	34	2	T08490	inorganic diphosph
580	5	3.7	37	2	B38230	hypothetical prote
581	5	3.7	44	1	SHNZS	inorganic diphosph
582	5	3.7	44	1	AH1966	small hydrophobic
583	5	3.7	52	2	C38230	photosystem I 4.8k
584	5	3.7	52	2	T00166	inorganic diphosph
585	5	3.7	52	2	A83679	hypothetical prote
586	5	3.7	53	2	S32547	hypothetical prote
587	5	3.7	53	2	T70030	cytochrome c552 -
588	5	3.7	53	2	C98053	Kalikkrein - mouse
589	5	3.7	54	2	E95069	hypothetical prote
590	5	3.7	56	1	G64051	hypothetical prote
591	5	3.7	57	2	E82733	ribosomal protein
592	5	3.7	58	2	A45824	hypothetical prote
593	5	3.7	59	2	F86779	hypothetical prote
594	5	3.7	59	2	D70252	hypothetical prote
595	5	3.7	60	2	E71571	hypothetical prote
596	5	3.7	61	2	C90971	hypothetical prote
597	5	3.7	61	2	H90901	probable lipoprote
598	5	3.7	61	2	E70537	probable lipoprote
599	5	3.7	62	1	NINJ1M	hypothetical prote
600	5	3.7	62	1	NINJ3M	short neurotoxin 1
601	5	3.7	62	1	S35098	trypsin inhibitor
602	5	3.7	62	2	G82476	trypsin inhibitor
603	5	3.7	62	2	AF0468	hypothetical prote
604	5	3.7	63	2	T03494	probable lipoprote
605	5	3.7	63	2	AP2194	hypothetical prote
606	5	3.7	64	2	T48412	hypothetical prote
607	5	3.7	64	2	E90503	hypothetical prote
608	5	3.7	65	2	A30478	NADH2 dehydrogenas
609	5	3.7	65	2	C35947	crocinamine 3 precu
610	5	3.7	65	2	A35947	crocinamine 1 precu
611	5	3.7	65	2	JC5324	mycotoxin a precurs
612	5	3.7	65	2	AB3381	heavy metal bindin
613	5	3.7	66	1	TIAC	trypsin inhibitor
614	5	3.7	614	5	RSBS29	ribosomal protein
615	5	3.7	615	5	S40196	ribosomal protein
616	5	3.7	616	5	E40361	virC-region hypoch
617	5	3.7	617	5	E97739	D-alanyl-D-alanine
618	5	3.7	618	5	AE0078	hypothetical prote
619	5	3.7	619	5	F70706	hypothetical prote
620	5	3.7	620	5	B54377	hypothetical prote
621	5	3.7	621	5	G75310	interleukin-1 rece
622	5	3.7	622	5	T07575	hypothetical prote
623	5	3.7	623	5	AF0282	ribosomal protein
624	5	3.7	624	5	E87267	transcription prote
625	5	3.7	625	5	JQ2195	transcription regu
626	5	3.7	626	5	S75780	hypothetical prote
627	5	3.7	627	5	AD1909	hypothetical prote
628	5	3.7	628	5	D87340	hypothetical prote
629	5	3.7	629	5	G87164	hypothetical prote
630	5	3.7	630	5	B82597	hypothetical prote
631	5	3.7	631	5	A36809	hypothetical prote
632	5	3.7	632	5	T42944	hypothetical prote
633	5	3.7	633	5	B90904	hypothetical prote
634	5	3.7	634	5	D81205	hypothetical prote
635	5	3.7	635	5	T23357	hypothetical prote
636	5	3.7	636	5	B43856	hypothetical prote
637	5	3.7	637	5	T03860	hypothetical prote
638	5	3.7	638	5	AC2038	hypothetical prote
639	5	3.7	639	5	D30502	hypothetical prote
640	5	3.7	640	5	S47158	Ig kappa chain V r
641	5	3.7	641	5	H91202	metallothionein II
642	5	3.7	642	5	D69010	hypothetical prote
643	5	3.7	643	5	T35577	hypothetical prote
644	5	3.7	644	5	S25700	hypothetical prote
645	5	3.7	645	5	G86731	hypothetical prote
646	5	3.7	646	5	H82126	exodeoxyribonuclea
647	5	3.7	647	5	E83139	hypothetical prote
648	5	3.7	648	5	T17589	exodeoxyribonuclea
649	5	3.7	649	5	S13840	hypothetical prote
650	5	3.7	650	5	S44988	hypothetical prote
651	5	3.7	651	5	T37952	hypothetical prote
652	5	3.7	652	5	AH2795	hypothetical prote
653	5	3.7	653	5	I46505	conserved hypochet
654	5	3.7	654	5	B90706	MHC class II RLA-D
655	5	3.7	655	5	S07352	hypothetical prote
656	5	3.7	656	5	E85556	hypothetical prote
657	5	3.7	657	5	AC2848	hypothetical prote
658	5	3.7	658	5	AF3159	hypothetical prote
659	5	3.7	659	5	B64791	hypothetical prote
660	5	3.7	660	5	B82422	conserved hypochet
661	5	3.7	661	5	S46930	teg22 protein - m
662	5	3.7	662	5	A27144	larval serum prote
663	5	3.7	663	5	S58175	acyl carrier prote
664	5	3.7	664	5	S45320	translation elonga
665	5	3.7	665	5	C90955	hypothetical prote
666	5	3.7	666	5	B86743	conserved hypochet
667	5	3.7	667	5	T38941	zinc finger protei
668	5	3.7	668	5	G58932	ribosomal protein
669	5	3.7	669	5	BVBXK1	MAK3 protein - ye
670	5	3.7	670	5	A82800	hypothetical prote
671	5	3.7	671	5	JQ1863	guanylate kinase-r
672	5	3.7	672	5	D72269	conserved hypochet
673	5	3.7	673	5	T33460	hypothetical prote
674	5	3.7	674	5	D97294	DNA-binding protei
675	5	3.7	675	5	H85803	hypothetical prote
676	5	3.7	676	5	S24236	hypothetical prote
677	5	3.7	677	5	AE2523	hypothetical prote
678	5	3.7	678	5	T42944	TCA3 protein - mou
679	5	3.7	679	5	I60330	variable region-al
680	5	3.7	680	5	A28762	Ig kappa chain V r
681	5	3.7	681	5	C86896	hypothetical prote
682	5	3.7	682	5	C69886	hypothetical prote
683	5	3.7	683	5	AC0406	hypothetical prote
684	5	3.7	684	5	AB2511	hypothetical prote
685	5	3.7	685	5	AB2511	neutrophysin 1 -
686	5	3.7	686	5	A60357	go

687	5	3.7	93	2	IS4417	MHC class II E-bet	760	5	3.7	107	2	S41542	membrane alanyl am
688	5	3.7	93	2	H97802	hypothetical prote	761	5	3.7	107	2	D83227	hypothetical prote
689	5	3.7	93	2	A89057	protein K09H1.5 I	762	5	3.7	107	2	F72571	hypothetical prote
690	5	3.7	93	2	D71257	hypothetical prote	763	5	3.7	107	2	E71213	hypothetical prote
691	5	3.7	94	2	E60748	MHC class II histo	764	5	3.7	108	2	F81982	hypothetical prote
692	5	3.7	94	2	E60748	MHC class II histo	765	5	3.7	108	2	F31565	hypothetical prote
693	5	3.7	94	2	E60748	MHC class II histo	766	5	3.7	108	2	AD1714	thioredoxin homolo
694	5	3.7	94	2	E60748	MHC class II histo	767	5	3.7	108	2	AD1343	thioredoxin homolo
695	5	3.7	94	2	A60748	MHC class II histo	768	5	3.7	109	2	B70037	hypothetical prote
696	5	3.7	94	2	H71888	dna-binding protei	769	5	3.7	109	2	F88021	hypothetical prote
697	5	3.7	94	2	C64624	histone-like DNA-b	770	5	3.7	109	2	C84386	hypothetical prote
698	5	3.7	94	2	T31022	conserved hypotnet	771	5	3.7	110	2	A71454	hypothetical prote
699	5	3.7	94	2	AB0522	probable secreted	772	5	3.7	110	2	B34501	hypothetical prote
700	5	3.7	94	2	AB1456	hypothetical prote	773	5	3.7	111	2	S75752	hypothetical prote
701	5	3.7	95	2	B87216	conserved membrane	774	5	3.7	112	2	I26317	hypothetical prote
702	5	3.7	95	2	G84225	hypothetical prote	775	5	3.7	112	2	H26317	Ig kappa chain V r
703	5	3.7	95	2	S07013	hypothetical prote	776	5	3.7	112	2	E26317	Ig kappa chain V r
704	5	3.7	95	2	B64133	hypothetical prote	777	5	3.7	112	2	G26317	Ig kappa chain V r
705	5	3.7	95	2	F71731	glutaredoxin 3 (gr	778	5	3.7	112	2	C26317	Ig kappa chain V r
706	5	3.7	95	2	F75438	glutaredoxin 3 (gr	779	5	3.7	112	2	B26317	Ig kappa chain V r
707	5	3.7	95	2	B45519	variant surface gl	780	5	3.7	112	2	A26317	Ig kappa chain V r
708	5	3.7	95	2	T26179	hypothetical prote	781	5	3.7	112	2	D26317	Ig kappa chain V r
709	5	3.7	96	2	H97835	integrator host r	782	5	3.7	112	2	F26317	Ig kappa chain V r
710	5	3.7	96	2	G69745	hypothetical prote	783	5	3.7	112	2	PL0275	Ig kappa chain V r
711	5	3.7	96	2	T03280	probable lipid tra	784	5	3.7	112	2	PL0274	Ig kappa chain V r
712	5	3.7	96	2	S24962	hypothetical prote	785	5	3.7	112	2	B49060	Ig kappa chain V r
713	5	3.7	96	2	AD0079	ecp1 protein - fun	786	5	3.7	112	2	T25554	Ig kappa chain V r
714	5	3.7	96	2	AH3461	hypothetical prote	787	5	3.7	112	2	G64953	hypothetical prote
715	5	3.7	97	2	C45681	hypothetical prote	788	5	3.7	112	2	A84369	hypothetical prote
716	5	3.7	97	2	C37202	hypothetical prote	789	5	3.7	112	2	B41871	hypothetical prote
717	5	3.7	98	1	DNH0NL	hypothetical 11k p	790	5	3.7	112	2	B95331	hypothetical prote
718	5	3.7	98	1	T14198	NADH2 dehydrogenas	791	5	3.7	112	2	AB1447	hypothetical prote
719	5	3.7	98	1	T11415	NADH2 dehydrogenas	792	5	3.7	112	2	AE1361	Protein gp10 (bact
720	5	3.7	98	1	QXBO4L	NADH2 dehydrogenas	793	5	3.7	113	2	B75632	Ig kappa chain V r
721	5	3.7	98	2	A58893	NADH2 dehydrogenas	794	5	3.7	113	2	AE0176	hypothetical prote
722	5	3.7	98	2	S41843	NADH2 dehydrogenas	795	5	3.7	113	2	S44750	conserved hypotnet
723	5	3.7	98	2	S26159	NADH2 dehydrogenas	796	5	3.7	113	2	AB1474	C06G4.3 protein -
724	5	3.7	98	2	T14025	NADH2 dehydrogenas	797	5	3.7	113	2	AF1112	PR5 beta-glucoside
725	5	3.7	98	2	T11865	NADH2 dehydrogenas	798	5	3.7	114	2	F97171	hypothetical prote
726	5	3.7	98	2	T11058	NADH2 dehydrogenas	799	5	3.7	114	2	E87690	hypothetical prote
727	5	3.7	98	2	T11371	NADH2 dehydrogenas	800	5	3.7	114	2	T50166	hypothetical prote
728	5	3.7	99	2	S49414	NADH2 dehydrogenas	801	5	3.7	114	2	AD2280	homolog to yeast g
729	5	3.7	99	2	JC2417	major carboxysome	802	5	3.7	115	2	T35387	hypothetical prote
730	5	3.7	99	2	C84541	monocyte chemotatr	803	5	3.7	115	2	E97457	hypothetical prote
731	5	3.7	100	2	AI0324	urease (EC 3.5.1.5	804	5	3.7	115	2	S09861	hypothetical prote
732	5	3.7	100	2	B43998	hypothetical prote	805	5	3.7	115	2	AH2014	transcription regu
733	5	3.7	100	2	A43998	hypothetical prote	806	5	3.7	115	2	R5HSH9	ribosomal protein
734	5	3.7	101	2	E83252	hypothetical prote	807	5	3.7	116	2	AD1189	NADH2 dehydrogenas
735	5	3.7	101	2	AB0506	probable transcrip	808	5	3.7	116	2	AE1547	phosphotransferase
736	5	3.7	102	2	S31417	urease (EC 3.5.1.5	809	5	3.7	116	2	D64681	nonspecific lipid
737	5	3.7	102	2	S48012	IP7 protein - phag	810	5	3.7	116	2	S50753	ribosomal protein
738	5	3.7	102	2	S39687	glutaredoxin 3 (lm	811	5	3.7	116	2	G90956	unknown protein en
739	5	3.7	102	2	H87678	hypothetical prote	812	5	3.7	116	2	D95280	hypothetical prote
740	5	3.7	102	2	C97733	methanogen chromos	813	5	3.7	116	2	T05517	hypothetical prote
741	5	3.7	103	2	G72577	hypothetical prote	814	5	3.7	116	2	S07933	abscisic acid-indu
742	5	3.7	103	2	T17940	hypothetical prote	815	5	3.7	116	2	F83302	hypothetical prote
743	5	3.7	103	2	B82626	probable prophage	816	5	3.7	116	2	C81779	hypothetical prote
744	5	3.7	103	2	AH0133	hypothetical prote	817	5	3.7	116	2	S76250	hypothetical prote
745	5	3.7	104	2	E72758	hypothetical prote	818	5	3.7	117	2	AC70027	conserved hypotnet
746	5	3.7	104	2	T35601	hypothetical prote	819	5	3.7	117	2	AG0844	multidrug-efflux t
747	5	3.7	104	2	T47010	hypothetical prote	820	5	3.7	117	2	AC2194	glutitol operon ac
748	5	3.7	105	2	JC2566	hypothetical prote	821	5	3.7	117	2	UQ1280	lipid transfer pro
749	5	3.7	105	2	S62940	b1a protein - Pseu	822	5	3.7	118	1	S57057	probable membrane
750	5	3.7	105	2	AB0628	conserved membrane	823	5	3.7	118	1	G65009	hypothetical prote
751	5	3.7	105	2	B72205	conserved hypotnet	824	5	3.7	119	2	AD4804	YbG protein - Esc
752	5	3.7	105	2	D72546	hypothetical prote	825	5	3.7	119	2	HSDU1A	histone H1a, sperm
753	5	3.7	105	2	C97846	hypothetical prote	826	5	3.7	119	2	H71351	probable ribosomal
754	5	3.7	106	2	AB1203	conserved hypotnet	827	5	3.7	120	1		
755	5	3.7	106	2	S32032	SP1 protein precu	828	5	3.7	120	1		
756	5	3.7	106	2	AC3648	flagellar motor sw	829	5	3.7	120	2		
757	5	3.7	107	2	PE0009	ornithine carbamoy	830	5	3.7	120	2		
758	5	3.7	107	2	H97859	hypothetical prote	831	5	3.7	121	1		
759	5	3.7	107	2	B71069	hypothetical prote	832	5	3.7	121	2		

833	5	3.7	121	2	H87664	glycine cleavage s	906	5	3.7	135	1	R5H032	ribosomal protein
834	5	3.7	121	2	B97446	hypothetical prote	907	5	3.7	135	1	R5M532	ribosomal protein
835	5	3.7	121	2	AC2664	ATP synthase, subu	908	5	3.7	135	1	R5R132	ribosomal protein
836	5	3.7	121	2	G75604	hypothetical prote	909	5	3.7	135	1	T40038	hypothetical prote
837	5	3.7	122	2	S47702	Yjnh protein - Esc	910	5	3.7	135	2	S35025	nolV protein - Rhi
838	5	3.7	122	2	T04554	hypothetical prote	911	5	3.7	135	2	H70516	hypothetical prote
839	5	3.7	122	2	E64837	ycv protein - Esc	912	5	3.7	135	2	F84678	hypothetical prote
840	5	3.7	122	2	B90760	hypothetical prote	913	5	3.7	135	2	C83051	hypothetical prote
841	5	3.7	122	2	H48523	hypothetical prote	914	5	3.7	135	2	G97172	hypothetical prote
842	5	3.7	122	2	T48851	hypothetical prote	915	5	3.7	136	2	F64107	ribonuclease P (EC
843	5	3.7	122	2	C97815	hydrogenase (Ec 1.	916	5	3.7	136	2	S28723	phosphoribosyl-AMP
844	5	3.7	123	1	H0DVFS	hypothetical prote	918	5	3.7	136	2	S34196	Alu RNA-binding pr
845	5	3.7	123	1	B64455	hypothetical prote	919	5	3.7	136	2	H81440	signal recognition
846	5	3.7	123	1	S56967	hypothetical prote	920	5	3.7	136	2	I53300	probable integral
847	5	3.7	123	2	S72389	hypothetical prote	921	5	3.7	137	2	H81452	Interleukin-1-beta
848	5	3.7	123	2	T20354	hypothetical prote	922	5	3.7	137	2	S22515	nucleoside diphosp
849	5	3.7	123	2	T49377	hypothetical prote	923	5	3.7	137	2	B87407	thlolin precursor,
850	5	3.7	123	2	AG0283	probable exported	924	5	3.7	137	2	G96017	ribosomal protein
851	5	3.7	123	2	C81441	hypothetical prote	925	5	3.7	138	2	T44080	conserved hypotet
852	5	3.7	124	2	A64946	hypothetical prote	926	5	3.7	138	2	I39641	potassium-transpor
853	5	3.7	124	2	G90947	hypothetical prote	927	5	3.7	138	2	S59086	toxin I - Actinoba
854	5	3.7	124	2	C85796	hypothetical prote	928	5	3.7	138	2	T04516	ribosomal protein
855	5	3.7	124	2	C85796	hypothetical prote	929	5	3.7	139	2	S60916	hypothetical prote
856	5	3.7	125	1	ZTBPT9	gene 50 protein -	930	5	3.7	139	2	H90442	probable membrane
857	5	3.7	125	1	C71669	30S ribosomal prot	931	5	3.7	139	2	A87442	hypothetical prote
858	5	3.7	125	2	H97822	synapsin I - mouse	932	5	3.7	140	2	D72680	Mct1/nudix family
859	5	3.7	125	2	A53692	hypothetical prote	933	5	3.7	140	2	D72680	hypothetical prote
860	5	3.7	125	2	AF1932	hypothetical prote	934	5	3.7	141	1	HABRM	hemoglobin alpha c
861	5	3.7	125	2	A86789	hypothetical prote	935	5	3.7	141	1	HABRT	hemoglobin alpha c
862	5	3.7	125	2	C83760	synapsin II - mus	936	5	3.7	141	1	HATGP	hemoglobin alpha-1
863	5	3.7	126	2	I61260	transcription acti	937	5	3.7	141	1	B71504	hemoglobin alpha-D
864	5	3.7	126	2	S65805	hypothetical prote	938	5	3.7	141	2	C81666	nucleoside-diphosp
865	5	3.7	126	2	E82404	hypothetical prote	939	5	3.7	141	2	A26539	hemoglobin alpha c
866	5	3.7	126	2	C87672	hypothetical prote	940	5	3.7	141	2	E75176	transcription regu
867	5	3.7	126	2	H87294	hypothetical prote	941	5	3.7	141	2	H71037	probable transcrip
868	5	3.7	126	2	G84215	hypothetical prote	942	5	3.7	141	2	T46972	transcription regu
869	5	3.7	126	2	VXECSE	preprotein translo	943	5	3.7	141	2	G70144	hypothetical prote
870	5	3.7	127	2	T36833	preprotein translo	944	5	3.7	141	2	T45687	hypothetical prote
871	5	3.7	127	2	H91241	preprotein translo	945	5	3.7	141	2	H90016	hypothetical prote
872	5	3.7	127	2	E86089	preprotein translo	946	5	3.7	142	2	A64074	hypothetical prote
873	5	3.7	127	2	A10456	preprotein translo	947	5	3.7	142	2	B82338	ribosomal protein
874	5	3.7	127	2	AC0934	preprotein translo	948	5	3.7	142	2	AB0255	probable phase pro
875	5	3.7	127	2	B89804	preprotein translo	949	5	3.7	142	2	S75218	hypothetical prote
876	5	3.7	128	1	AZPSCD	conserved hypotet	950	5	3.7	142	2	H72520	hypothetical prote
877	5	3.7	128	2	G71663	azurin - Pseudomon	951	5	3.7	142	2	C81359	hypothetical prote
878	5	3.7	128	2	E90328	cytochrome-c biosy	952	5	3.7	143	1	S04230	H+-transporting tw
879	5	3.7	128	2	T50410	very hypotetical	953	5	3.7	143	2	S34069	H+-transporting tw
880	5	3.7	128	2	S72600	hypothetical prote	954	5	3.7	143	2	F87328	ribosomal protein
881	5	3.7	129	2	S20611	calpastatin - rat	955	5	3.7	143	2	A12132	hypothetical prote
882	5	3.7	129	2	F81310	hypothetical prote	956	5	3.7	143	2	F64634	hypothetical prote
883	5	3.7	129	2	T19263	hypothetical prote	957	5	3.7	143	2	G71880	hypothetical prote
884	5	3.7	129	2	S73598	Me399 homolog D02-	958	5	3.7	143	2	T23735	hypothetical prote
885	5	3.7	129	2	B83350	hypothetical prote	959	5	3.7	143	2	B86857	conserved hypotet
886	5	3.7	129	2	E82143	conserved hypotet	960	5	3.7	144	2	H86567	nucleoside-2'-P kin
887	5	3.7	129	2	T45057	hypothetical prote	961	5	3.7	144	2	G72056	nucleoside diphosp
888	5	3.7	130	2	S40321	Ig kappa chain - h	962	5	3.7	144	2	A81033	hypothetical prote
889	5	3.7	130	2	F71035	hypothetical prote	963	5	3.7	144	2	B81977	hypothetical prote
890	5	3.7	131	2	T44990	cransposase (impor	964	5	3.7	144	2	D69936	conserved hypotet
891	5	3.7	131	2	G69959	hypothetical prote	965	5	3.7	144	2	E72647	hypothetical prote
892	5	3.7	131	2	A81137	Tspp-related prote	966	5	3.7	144	2	B24500	hypothetical prote
893	5	3.7	131	2	I46955	retinol-binding pr	967	5	3.7	144	2	T41544	hypothetical prote
894	5	3.7	132	2	T37861	hypothetical prote	968	5	3.7	145	2	G72451	hypothetical prote
895	5	3.7	132	2	AH2171	hypothetical prote	969	5	3.7	145	2	AD2345	hypothetical prote
896	5	3.7	133	1	CYRGG	gamma-crystallin I	970	5	3.7	145	2	T20956	hypothetical prote
897	5	3.7	133	2	B75526	hypothetical prote	971	5	3.7	145	2	T33037	hypothetical prote
898	5	3.7	133	2	G81710	hypothetical prote	972	5	3.7	145	2	G71158	hypothetical prote
899	5	3.7	133	2	F72471	hypothetical prote	973	5	3.7	146	1	G69131	ribosomal protein
900	5	3.7	133	2	B49530	hypothetical prote	974	5	3.7	146	2	D64733	prelipin peptidase
901	5	3.7	133	2	A84221	vascular endotheli	975	5	3.7	146	2	H90642	prelipin peptidase
902	5	3.7	134	2	A30563	riboflavin synthas	976	5	3.7	146	2	H85493	hypothetical prote
903	5	3.7	134	2	D70518	T-cell receptor be	977	5	3.7	146	2	D70766	hypothetical prote
904	5	3.7	134	2	J01403	acidic seminal flu	978	5	3.7	146	2		
905	5	3.7	134	2	AH3177	conserved hypotet	978	5	3.7	146	2		

979 5 3.7 146 2 AB0623
 980 5 3.7 146 2 T36436
 981 5 3.7 146 2 A12376
 982 5 3.7 147 2 S21648
 983 5 3.7 147 2 A26697
 984 5 3.7 147 2 S68351
 985 5 3.7 147 2 E83381
 986 5 3.7 147 2 H82172
 987 5 3.7 147 2 B97581
 988 5 3.7 147 2 AH2801
 989 5 3.7 147 2 AD3381
 990 5 3.7 148 2 AF1081
 991 5 3.7 148 2 AG1438
 992 5 3.7 148 2 E71822
 993 5 3.7 148 2 T05143
 994 5 3.7 148 2 H72781
 995 5 3.7 148 2 A91290
 996 5 3.7 148 2 F86131
 997 5 3.7 148 2 D49530
 998 5 3.7 149 2 R5BS7F
 999 5 3.7 149 2 A69220
 1000 5 3.7 149 2 T01433

probable bacterioph
 probable gntf-fami
 hypothetical prote
 T-cell receptor al
 echinoidin - sea u
 heme d1 synthesis
 probable transcrip
 heat shock protein
 hypothetical prote
 mult like protein
 heeb protein [limp
 50S ribosomal prot
 50S ribosomal prot
 hypothetical prote
 hypothetical prote
 hypothetical prote
 hypothetical prote
 16K vascular endot
 ribosomal protein
 conserved hypotet
 alternative respir

ALIGNMENTS

RESULT 1
 A13337
 hypothetical protein BME10687 [Imported] - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
 C:Accession: A13337
 R:Deleucchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, P.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
 A:Reference number: AD2552; PMID:11756688
 A:Accession: A13337
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-100 <KUR>
 A:Cross-references: GB:AE008917; PIDN:AL51868.1; PID:g17982618; GSPDB:GNO0190
 C:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BME10687
 A:Map position: 1

Query Match
 Best Local Similarity 100.0%; Score 8; DB 2; Length 100;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AFTLASCA 21
 |||||
 Db 71 AFTLASCA 78

RESULT 2
 S37595
 mucin J7U110 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Oct-1999
 C:Accession: S37595; S35048
 R:Audert, J.
 submitted to the EMBL Data Library, September 1993
 A:Reference number: S37593
 A:Accession: S37595
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-330 <AUB>
 A:Cross-references: EMBL:X74956; NID:9407052; PIDN:CAA52911.1; PID:g407053
 R:Dufosse, J.; Porchet, N.; Audie, J.P.; Guyonnet Duperat, V.; Laine, A.; van-Seuning, Biochem. J. 293, 329-337, 1993

A:Title: Degenerate 87-base-pair tandem repeats create hydrophilic/hydrophobic altern
 A:Reference number: S35047; MUID:93343858; PMID:7916618
 A:Accession: S35048
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-330 <DUF>

Query Match
 Best Local Similarity 100.0%; Score 8; DB 2; Length 330;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 LTELTTAA 98
 |||||
 Db 26 LTELTTAA 33

RESULT 3

S35047
 mucin J7U7 - human
 C:Species: Homo sapiens (man)
 C>Date: 10-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 08-Oct-1999
 C:Accession: S35047
 R:Dufosse, J.; Porchet, N.; Audie, J.P.; Guyonnet Duperat, V.; Laine, A.; van-Seuning, Biochem. J. 293, 329-337, 1993
 A:Title: Degenerate 87-base-pair tandem repeats create hydrophilic/hydrophobic altern
 A:Reference number: S35047; MUID:93343858; PMID:7916618
 A:Accession: S35047
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-543 <DUF>
 A:Cross-references: EMBL:X74370; NID:9407081; PIDN:CAA52408.1; PID:g407082
 A:Note: the authors translated the codon AAA for residue 63 as Gln and Cgg for residu
 A:Note: the amino acid sequence from Fig. 2 is inconsistent with the nucleotide sequ

Query Match
 Best Local Similarity 100.0%; Score 8; DB 2; Length 543;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 LTELTTAA 98
 |||||
 Db 174 LTELTTAA 181

RESULT 4

S35049
 mucin JER57 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 10-Dec-1993 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: S35049; S37594
 R:Dufosse, J.; Porchet, N.; Audie, J.P.; Guyonnet Duperat, V.; Laine, A.; van-Seuning, Biochem. J. 293, 329-337, 1993
 A:Title: Degenerate 87-base-pair tandem repeats create hydrophilic/hydrophobic altern
 A:Reference number: S35047; MUID:93343858; PMID:7916618
 A:Accession: S35049
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-610 <DUF>
 R:Audert, J.
 submitted to the EMBL Data Library, September 1993
 A:Reference number: S37593
 A:Accession: S37594
 A:Molecule type: mRNA
 A:Residues: 1-20, 'w', 22-610 <AUB>
 A:Cross-references: EMBL:X74955

Query Match
 Best Local Similarity 100.0%; Score 8; DB 2; Length 610;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 LTELTTAA 98
 |||||
 Db 221 LTELTTAA 228

```
RESULT 5
T45025
mucin MUC5B, tracheobronchial [imported] - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T45025
R:Deasy, J.L.; Guyonnet-Duperrat, V.; Porchet, N.; Aubert, J.P.; Laine, A.
J. Biol. Chem. 272, 3168-3178, 1997
A:Title: Human mucin gene MUC5B, the 10.7 kb large central exon encodes various alternat
A:Reference number: 222899; MUID:9716151; PMID:9013550
A:Accession: T45025
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3570 <DES>
A:Cross-references: EMBL:272496; NID:g1834502; PIND:CAA96577.1; PID:g1834503
C:Genetics:
A:Gene: MUC5B

Query Match
Best Local Similarity 100.0%; Score 8; DB 2; Length 3570;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 91 LTELTTAA 98
|||||
DB 3008 LTELTTAA 3015

RESULT 6
G97259
uncharacterized protein, possibly involved in thiamine biosynthesis [imported] - Clostr
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: G97259
R:Molling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A:Reference number: A96900; MUID:2135935; PMID:21359325
A:Accession: G97259
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-64 <KUR>
A:Cross-references: GB:AE001437; PIND:AAK80866.1; PID:g15025974; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC2924

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 KLSSEK 70
|||||
DB 48 KLSSEK 54

RESULT 7
AE3445
ATP synthase BMEI1547 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AE3445
R:Delvecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A:Reference number: AD3252; PMID:11756688
A:Accession: AE3445
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-132 <KUR>
```

```
A:Cross-references: GB:AE008917; PIND:AAL52728.1; PID:g17983558; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI1547
A:Map position: 1

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 132;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 AMKLSE 68
|||||
DB 62 AMKLSE 68

RESULT 8
VCEVUR
coat protein gp37 - avian sarcoma virus UR2
N:Alternate names: env protein gp37
C:Species: avian sarcoma virus UR2
A:Note: host Gallus gallus (chicken)
C:Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 12-May-1994
C:Accession: A03998
R:Neckameyer, W.S.; Wang, L.H.
J. Virol. 53, 879-884, 1985
A:Title: Nucleotide sequence of avian sarcoma virus UR2 and comparison of its transfo
A:Reference number: A00635; MUID:65136034; PMID:2983097
A:Accession: A03998
A:Molecule type: genomic RNA
A:Residues: 1-174 <NEC>
C:Genetics:
A:Gene: env
C:Superfamily: type C retrovirus env polyprotein
C:Keywords: coat protein; polyprotein

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 174;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39
|||||
DB 21 ANLTSL 27

RESULT 9
AH3296
dUTP diphosphatase (EC 3.6.1.23) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 03-Jun-2002
C:Accession: AH3296
R:Delvecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella meli
A:Reference number: AD3252; PMID:11756688
A:Accession: AH3296
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-174 <KUR>
A:Cross-references: GB:AE008917; PIND:AAL51539.1; PID:g17982257; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI058
A:Map position: 1
C:Superfamily: dUTP pyrophosphatase
C:Keywords: hydrolase

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 174;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 71 AKISETA 77
|||||
```

Db 156 AKISETA 162

RESULT 10
B82380
purine-binding chemotaxis protein CheW VCA1094 [imported] - Vibrio cholerae (strain N169)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: B82380
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
chadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: B82380
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-175 <HEI>
A:Cross-references: GB:AE004434; GB:AE003853; NID:9658531; PIDN:AAF96986.1; GSPDB:GN001
C:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA1094
A:Map position: 2

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 175;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 32 SANLITS 38
3 SANLITS 9

RESULT 11
AH3404
hypothetical protein BME1122 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AH3404
R:DeIvecchio, V.G.; Kaputral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:1156688
A:Accession: AH3404
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-186 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL52403.1; PID:G17983204; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME1122
A:Map position: 1

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 186;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 106 EVAOKIV 112
16 EVAOKIV 22

RESULT 12
CA8613
env polyprotein TM - avian myeloblastosis virus
C:Species: avian myeloblastosis virus
C>Date: 21-Jan-1994 #sequence_revision 25-Apr-1997 #text_change 05-May-2000
C:Accession: CA8613
R:Joliet, V.; Borroughs, K.; Lasserre, F.; Crochet, J.; Dambrie, G.; Smith, R.E.; Perbal
Virology 195, 812-819, 1993
A:Title: Pathogenic potential of myeloblastosis-associated virus: implication of env pro

A:Reference number: A48613; MUID:93331743; PMID:8393249
A:Contents: MAY1(N)/2
A:Accession: CA8613
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-198 <JOL>
A>Note: sequence extracted from NCBI backbone (NCBI:135489)
C:Superfamily: type C retrovirus env polyprotein

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 198;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 33 ANLITSL 39
51 ANLITSL 57

RESULT 13
D96842
hypothetical protein F23A5.30 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96842
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maitl, R.; Marzla
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96842
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-213 <STO>
A:Cross-references: GB:AE005173; NID:96503306; PIDN:AAF14682.1; GSPDB:GN00141
C:Genetics:
A:Gene: F23A5.30
A:Map position: 1

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 213;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 83 DKAPEAV 89
44 DKAPEAV 50

RESULT 14
VCFV37
coat protein gp37 - Rous sarcoma virus (fragment)
C:Species: Rous sarcoma virus
C>Date: 18-Dec-1981 #sequence_revision 19-Feb-1984 #text_change 24-Sep-1999
C:Accession: B38017; B38018; A03397
R:Czerwikofsky, A.P.; Levinson, A.D.; Varmus, H.E.; Bishop, J.M.; Fischer, E.; Goodma
Nature 301, 736-738, 1983
A:Title: Corrections to the nucleotide sequence of the src gene of Rous sarcoma virus
A:Reference number: A38017; MUID:83141780; PMID:6298633
A:Accession: B38017
A:Molecule type: DNA
A:Residues: 1-246 <CZE>
A:Cross-references: GB:I29199; GB:J02018; GB:J02026; GB:J02352; GB:K01194; GB:K01195;
R:Flakey, T.; Feldman, R.A.; Hanafusa, H.
J. Virol. 44, 1-11, 1982
A:Title: DNA sequence of the viral and cellular src gene of chickens: I. Complete nuc
A:Reference number: A38018; MUID:83059858; PMID:6292477
A:Accession: B38018

A:Molecule type: DNA
 A:Residues: 1-20, 'I', '29', 'A', '31-42', 'V', '44-78', 'K', '80-144', 'H', '146-168', 'P', '170-211'
 A:Cross-references: GB:K00928; NID:9210187; PIDN:AAA42564.1; PID:9210188
 A:Experimental source: strain Schmidt-Rupplin
 C:Genetics:

A:Gene: env
 C:Superfamily: type C retrovirus env polyprotein

Query Match 5.2%; Score 7; DB 1; Length 246;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39
 |||||||
 Db 92 ANLTSL 98

RESULT 15

env polyprotein ev21 - avian endogenous virus ev21 (fragment)

C:Species: avian endogenous virus ev21
 C:Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 18-Feb-2000
 C:Accession: I50661, S23734

R:Levin, I.; Smith, E.J.
 Poult. Sci. 70, 1948-1956, 1991
 A:Title: Association of a chicken repetitive element with the endogenous virus-21 slow-f
 A:Reference number: I50661; MUID:92141069; PMID:1685775
 A:Accession: I50661

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-246 <LEV>

A:Cross-references: EMBL:X54094; NID:963547; PIDN:CAA38028.1; PID:9388547
 A:Note: submitted to the EMBL Data Library, July 1990
 C:Genetics:

A:Gene: env

C:Superfamily: type C retrovirus env polyprotein

C:Keywords: polyprotein

Query Match 5.2%; Score 7; DB 2; Length 246;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39
 |||||||
 Db 92 ANLTSL 98

RESULT 16

env polyprotein - avian endogenous virus RAV-0 (fragment)

N:Contains: 37K glycoprotein; 85K glycoprotein
 C:Species: avian endogenous virus RAV-0
 C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 20-Sep-1999

R:Trihomenko, A.T.; Lomovskaya, O.L.
 Virus Genes 3, 251-258, 1990
 A:Title: Avian endogenous provirus (ev-3) env gene sequencing: implication for pathogeni
 A:Reference number: A60398; MUID:90266494; PMID:2161159
 A:Accession: A60398

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-246 <TIK>
 A:Cross-references: EMBL:X07818; NID:961462; PIDN:CAA30677.1; PID:9833162
 C:Genetics:

A:Gene: env
 C:Superfamily: type C retrovirus env polyprotein

C:Keywords: glycoprotein; polyprotein

Query Match 5.2%; Score 7; DB 2; Length 246;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39

Db 92 ANLTSL 98
 |||||||

RESULT 17

env polyprotein - avian retrovirus RPL30 (fragment)

C:Species: avian retrovirus RPL30
 C:Date: 11-Feb-1993 #sequence_revision 30-Sep-1993 #text_change 16-Feb-1997
 C:Accession: A43362

R:Jia, R.; Mayer, B.J.; Hanafusa, T.; Hanafusa, H.
 J. Virol. 66, 5975-5987, 1992

A:Title: A novel oncogene, v-rk, encoding a truncated receptor tyrosine kinase is tr
 A:Reference number: A43362; MUID:92407992; PMID:1527848
 A:Accession: A43362

A:Molecule type: genomic RNA

A:Residues: 1-257 <JIA>

A:Cross-references: GB:M92847

C:Genetics:

A:Gene: env
 C:Superfamily: type C retrovirus env polyprotein

C:Keywords: coat protein; polyprotein; transmembrane protein

Query Match 5.2%; Score 7; DB 2; Length 257;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39
 |||||||
 Db 129 ANLTSL 135

RESULT 18

hypothetical protein jhp0823 - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori
 A:Variety: strain J99
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 21-Jan-2000

C:Accession: A71885

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.
 ; Ives, C.; Gibson, R.; Meberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F
 Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p
 A:Reference number: A71800; MUID:9120557; PMID:9923682

A:Accession: A71885

A:Status: preliminary

A:Molecule type: DNA

A:Cross-references: GB:AE001511; GB:AE001439; NID:94155382; PIDN:AAD06395.1; PID:9415

A:Experimental source: strain J99

C:Genetics:

A:Gene: jhp0823
 C:Superfamily: retinol dehydrogenase; short-chain alcohol dehydrogenase homology
 F.18-173/Domain: short-chain alcohol dehydrogenase homology <SMD>

Query Match 5.2%; Score 7; DB 2; Length 271;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 106 EVAOKIV 112
 |||||||
 Db 217 EVAOKIV 223

RESULT 19

hypothetical protein F17A9.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T31710

R:Woldmann, P.; Sammons, L.; Rohlfing, T.; Gilliam, B.
 submitted to the EMBL Data Library, July 1997

A:Description: The sequence of C. elegans cosmid F17A9.

A:Reference number: Z21072
A:Accession: T31710.
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-277 <MOH>
A:Cross-references: EMBL:AF016417, PIDD:AA65278.1; GSPDB:GN00023; CESP:F17A9.1
A:Experimental source: strain Bristol N2; clone F17A9
C:Genetics:
A:Gene: CESP:F17A9.1
A:Map position: 5
A:Introns: 66/2; 94/2; 120/2; 198/3

Query Match
Best Local Similarity 5.2%; Score 7; DB 2; Length 277;
C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 15-Sep-2000
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 90 SLTSLT 96
|||||||
DB 184 SLTSLT 190

RESULT 20
E64216
hypothetical protein MG149 - Mycoplasma genitalium
C:Species: Mycoplasma genitalium
C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 15-Sep-2000
A:Accession: E64216
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;
M.; Fuhmann, J.; Nguyen, D.; Uterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.
; C.A.; Venter, J.C
Science 270, 397-403, 1995
A:Title: The minimal gene complement of Mycoplasma genitalium.
A:Reference number: A64200; MUID:96026346; PMID:756993
A:Accession: E64216
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-281 <TIGR>
A:Cross-references: GB:039694; GB:LA3967; NID:91045822; PID:91045832; TIGR:MG149
A:Experimental source: strain G-37
C:Genetics:
A:Genetic code: SGC3
C:Superfamily: Mycoplasma pneumoniae probable lipoprotein VPSPT7_orf320

Query Match
Best Local Similarity 5.2%; Score 7; DB 2; Length 281;
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 LASCAS 23
|||||||
DB 21 LASCAS 27

RESULT 21
T09452
vdiC protein - Helicobacter pylori (strain 60190)
C:Species: Helicobacter pylori
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
A:Accession: T09452
R:Caio, P.; Cover, T.L.
J. Bacteriol. 179, 2852-2856, 1997
A:Title: High-level genetic diversity in the vapD chromosomal region of Helicobacter pylori
A:Reference number: Z16675; MUID:97284485; PMID:9139899
A:Accession: T09452
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-284 <CAO>
A:Cross-references: EMBL:U94318; NID:92072451; PIDD:AA45243.1; PID:92072454
C:Genetics:
A:Gene: vdiC
C:Superfamily: retinol dehydrogenase; short-chain alcohol dehydrogenase homology
F:29-184/Domain: short-chain alcohol dehydrogenase homology <SAD>

Query Match 5.2%; Score 7; DB 2; Length 284;

Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 106 EVAOKIV 112
|||||||
DB 230 EVAOKIV 236

RESULT 22
D82112
Lyra protein VC2145 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
A:Accession: D82112
R:Heidelberg, J.F.; Eissen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.
; Chardson, D.; Ermolaeva, M.D.; Vamathavan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: D82112
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-302 <HEI>
A:Cross-references: GB:AE004287; GB:AE003852; NID:9656689; PIDD:AAE95290.1; GSPDB:GN
C:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2145
A:Map position: 1
C:Superfamily: Escherichia coli hypothetical protein b2431

Query Match
Best Local Similarity 5.2%; Score 7; DB 2; Length 302;
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 110 KIVRSL 116
|||||||
DB 227 KIVRSL 233

RESULT 23
T23674
hypothetical protein M02B1.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
A:Accession: T23674
R:Lighting, J.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19780
A:Accession: T23674
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-320 <WIL>
A:Cross-references: EMBL:Z81102; PIDD:CA803202.1; GSPDB:GN00022; CESP:M02B1.3
A:Experimental source: clone M02B1
C:Genetics:
A:Gene: CESP:M02B1.3
A:Map position: 4
A:Introns: 35/1; 100/3; 152/2; 252/2; 291/1

Query Match
Best Local Similarity 5.2%; Score 7; DB 2; Length 320;
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 NLTTSLI 40
|||||||
DB 91 NLTTSLI 97

RESULT 24
A64054
selenophosphate synthase - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae

C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 19-May-2000
 C/Accession: A64054
 R:Flatschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M. Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: A64054
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-323 <TIGR>
 A:Cross-references: GB:U32705; GB:L42023; NID:91573156; PIDN:AAC21869.1; PID:91573160; TIGR:G000000000
 C:Genetics:
 A:Start codon: GTG
 C:Superfamily: hydrogenase expression/formation protein hylp
 C:Keywords: selenocysteine biosynthesis

Query Match 5.2%; Score 7; DB 2; Length 323;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 106 EVAOKIV 112
 |||||
 Db 92 EVAOKIV 98

RESULT 25
 JC7183
 cathepsin Q (EC 3.4.22.-) precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 24-Oct-2000
 C/Accession: JC7183
 R:Sol-Church, K.; French, J.; Mason, R.W. Biochem. Biophys. Res. Commun. 267, 791-795, 2000
 A:Title: Cathepsin Q, a novel lysosomal cysteine protease highly expressed in placenta. A:Reference number: JC7183; MUID:20139729; PMID:10673370
 A:Accession: JC7183
 A:Molecule type: mRNA
 A:Residues: 1-343 <SOLO>
 A:Cross-references: GB:AF187323; NID:96010770; PIDN:AAFO1247.1; PID:96010771
 A:Experimental source: placenta
 C:Superfamily: papain
 C:Keywords: cysteine proteinase; glycoprotein; hydrolase; placenta
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-125/Domain: propeptide #status predicted <PRO>
 F:126-343/Product: cathepsin Q #status predicted <MAT>
 F:149,286,310/Active site: Cys, His, Asn #status predicted
 F:228,298/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.2%; Score 7; DB 2; Length 343;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 27 NPKNSA 33
 |||||
 Db 225 NPKNSA 231

RESULT 26
 C64082
 ATP-binding protein homolog HI0621 - Haemophilus influenzae (strain Rd KW20)
 C:Species: Haemophilus influenzae
 C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 02-Feb-2001
 C/Accession: C64082
 R:Flatschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M. Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. A:Reference number: A64000; MUID:95350630; PMID:7542800

A:Accession: C64082
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-345 <TIGR>
 A:Cross-references: GB:U32744; GB:L42023; NID:91573608; PIDN:AAC2280.1; PID:91573615
 A:Note: named as homolog to a protein from Escherichia coli
 C:Superfamily: probable ABC-type transport protein abc; ATP-binding cassette homology
 C:Keywords: ATP; nucleotide binding; P-loop
 F:21-217/Domain: ATP-binding cassette homology <ABC>
 F:38-45/Region: nucleotide-binding motif A (P-loop)

Query Match 5.2%; Score 7; DB 2; Length 345;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 SESRAKI 73
 |||||
 Db 111 SESRAKI 117

RESULT 27
 AB3099
 hypothetical protein rspe [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
 C:Species: Agrobacterium tumefaciens
 C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
 C/Accession: AB3099
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woerage, G.; Gallet, W.; Grant, C.; Guenther, D.; Kulyavin, T.; Levy, R.; Li, M.; McCl; Karp, P.; Romero, P.; Zhang, S. Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58. A:Reference number: AB3577; PMID:11743193
 A:Accession: AB3099
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-434 <KUR>
 A:Cross-references: GB:AE008689; PIDN:AAL45208.1; PID:917742888; GSPDB:GN00187
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: rspe
 A:Map position: linear chromosome

Query Match 5.2%; Score 7; DB 2; Length 434;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 71 AKISETA 77
 |||||
 Db 236 AKISETA 242

RESULT 28
 G98187
 rhizobioicin secretion protein rspe (AF141932) [imported] - Agrobacterium tumefaciens
 C:Species: Agrobacterium tumefaciens
 C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
 C/Accession: G98187
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Gold A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
 A:Reference number: A97359; PMID:11743194
 A:Accession: G98187
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-436 <KUR>
 A:Cross-references: GB:AE007870; PIDN:AAK99025.1; PID:915158817; GSPDB:GN00170
 C:Genetics:
 A:Gene: AGR_L_906
 A:Map position: linear chromosome

Query Match 5.2%; Score 7; DB 2; Length 436;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 71 AKISETA 77
 |||||||
 Db 238 AKISETA 244

RESULT 29
 T11084
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Falco peregrinus mitochondrion
 C:Species: mitochondrion Falco peregrinus
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
 C:Accession: T11084
 R:Mindell, D.P.; Sorenson, M.D.; Dimcheff, D.E.
 Proc. Natl. Acad. Sci. U.S.A. 95, 10693-10697, 1998
 A:Title: Multiple independent origins of mitochondrial gene order in birds.
 A:Reference number: Z17242
 A:Accession: T11084
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-459 <MIN>
 A:Cross-references: EMBL:AF090338; NID:94894462; PID:94894472; PIDN:PAD32501.1
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGCI
 C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
 C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 5.2%; Score 7; DB 2; Length 459;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 TATLMTA 14
 |||||||
 Db 398 TATLMTA 404

RESULT 30
 DCBYO
 ornithine decarboxylase (EC 4.1.1.17) - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YKJ184w
 C:Species: Saccharomyces cerevisiae
 C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 21-Jul-2000
 C:Accession: A28437; S34686; S38016
 R:Fonzi, W.A.; Sypherd, P.S.
 J. Biol. Chem. 267, 10127-10133, 1992
 A:Title: The gene and the primary structure of ornithine decarboxylase from Saccharomyces
 A:Reference number: A28437; MUID:87280032; PMID:3038869
 A:Accession: A28437
 A:Molecule type: DNA
 A:Residues: 1-466 <FON>
 A:Cross-references: GB:J02777; NID:9172069; PIDN:AAA34829.1; PID:9172070
 R:Wiemann, S.; Voss, H.; Schwaggar, C.; Rupp, T.; Stegemann, J.; Zimmermann, J.; Grothues
 submitted to the EMBL Data Library, July 1993
 A:Description: Sequencing and analysis of 51.5 kbpases on the left arm of chromosome X
 A:Reference number: S34679
 A:Accession: S34686
 A:Molecule type: DNA
 A:Residues: 1-466 <WIE>
 A:Cross-references: EMBL:X74151; NID:9450365; PIDN:CAA52254.1; PID:9395241
 A:Experimental source: strain S288C
 R:Wiemann, S.; Voss, H.; Schwaggar, C.; Rupp, T.; Grothues, D.; Sensen, C.; Stegemann, J.
 submitted to the Protein Sequence Database, March 1994
 A:Reference number: S37825
 A:Accession: S38016
 A:Molecule type: DNA
 A:Residues: 1-466 <WIE>
 A:Cross-references: EMBL:Z28184; NID:9486324; PIDN:CAA62027.1; PID:9486325; GSPDB:GN0002
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGD:SPB1; ORD1; MIPS:YKJ184w

A:Cross-references: SGD:S0001667; MIPS:YKJ184w
 A:Map position: 11L
 C:Superfamily: ornithine decarboxylase
 C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis
 F:116/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted
 F:411/Active site: Cys (shared with dimeric partner) #status predicted

Query Match 5.2%; Score 7; DB 1; Length 466;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 13 TATLMTA 19
 |||||||
 Db 326 TATLMTA 332

RESULT 31
 D70154
 replication initiation protein dnaA - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 26-Aug-1999
 C:Accession: D70154
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
 son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
 Nature 390, 580-586, 1997
 A:Authors: Smith, H.O.; Venter, J.C.
 A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
 A:Reference number: A70100; MUID:98065943; PMID:9403685
 A:Accession: D70154
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-486 <KLE>
 A:Cross-references: GB:AE001149; GB:AE000783; NID:92688348; PIDN:AAB91515.1; PID:9268
 A:Experimental source: strain B31
 C:Superfamily: replication initiation protein dnaA

Query Match 5.2%; Score 7; DB 2; Length 486;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 KILYVTA 9
 |||||||
 Db 208 KILYVTA 214

RESULT 32
 F69001
 hypothetical protein MTH101 - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Species: Methanobacterium thermoautotrophicum
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: F69001
 R:Smith, D.R.; Doucellette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T
 K.L.S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7153-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
 A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: F69001
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-499 <MTH>
 A:Cross-references: GB:AE000800; GB:AE000666; NID:92621130; PIDN:AAB84600.1; PID:9262
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH101

Query Match 5.2%; Score 7; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 VTATLMT 13
 |||||||

DB 454 VTATLMT 460

RESULT 33

S35341

Kettin - fruit fly (*Drosophila melanogaster*) (fragment)

C:Species: *Drosophila melanogaster*

C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 16-Feb-1997

C:Accession: S35341; S39405

R:Label: A.; Label: S.; Gastei, M.; Ferguson, C.; Barlow, D.P.; Leonard, K.; Bullard, E.

EMBO J. 12, 2863-2871, 1993

A:Title: Kettin, a large modular protein in the Z-disc of insect muscles.

A:Reference number: S35341; MUID:9332776; PMID:8335002

A:Accession: S35341

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-524 <LAK>

A:Cross-references: EMBL:X72709

R:Label: S.

Submitted to the EMBL Data Library, March 1993

A:Reference number: S39405

A:Accession: S39405

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 3-524, 'G' <LAB>

A:Cross-references: EMBL:X72709

C:Genetics:

A:Gene: FlyBase:Ket

A:Cross-references: FlyBase:Frgn0010396

Query Match	5.2%;	Score 7;	DB 2;	Length 524;
Best Local Similarity	100.0%;	Pred. No. 39;		
Matches 7;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 101 PNARTEV 107

DB 518 PNARTEV 524

RESULT 34

H48613

env polyprotein precursor - myeloblastosis-associated virus (strain MAV-2(O)p9) (fragment)

C:Species: myeloblastosis-associated virus

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 26-Feb-1998

C:Accession: H48613

R:Joliot, V.; Borouhs, K.; Lasserre, F.; Crochet, J.; Dambrine, G.; Smith, R.E.; Perbat

Virology 195, 812-819, 1993

A:Title: Pathogenic potential of myeloblastosis-associated virus: implication of env pro

A:Reference number: A48613; MUID:9331743; PMID:8393249

A:Accession: H48613

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-594 <JOL>

C:Superfamily: type C retrovirus env polyprotein

Query Match	5.2%;	Score 7;	DB 2;	Length 594;
Best Local Similarity	100.0%;	Pred. No. 43;		
Matches 7;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 33 ANLTSL 39

DB 447 ANLTSL 453

RESULT 35

TVPSA

env-sea polyprotein - avian erythroblastosis virus (strain S13) (fragment)

N:Contains: protein-tyrosine kinase (EC 2.7.1.112) sea

C:Species: avian erythroblastosis virus

C>Date: 30-Jun-1991 #sequence_revision 05-May-1995 #text_change 18-Feb-2000

C:Accession: A33902; B33902

R:Smith, D.R.; Vogt, P.K.; Hayman, M.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 5291-5295, 1989

A:Title: The v-sea oncogene of avian erythroblastosis retrovirus S13: another member
A:Reference number: A33902; MUID:89315783; PMID:2546151
A:Accession: A33902
A:Molecule type: DNA
A:Residues: 1-596 <SMI>
A:Cross-references: EMBL:M25158
C:Genetics:

A:Gene: env-sea

C:Superfamily: avian erythroblastosis virus env-sea polyprotein; protein kinase homol

C:Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; t

F;1-226/Region: env polyprotein gene-derived

F;227-596/Region: protein-tyrosine kinase sea gene-derived

F;284-552/Domain: protein kinase homology <KIN>

F;292-300/Region: protein kinase ATP-binding motif

F;318/Active site: Lys #status predicted

Query Match	5.2%;	Score 7;	DB 1;	Length 596;
Best Local Similarity	100.0%;	Pred. No. 44;		
Matches 7;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 33 ANLTSL 39

DB 79 ANLTSL 85

RESULT 36

I48613

env polyprotein precursor - myeloblastosis-associated virus (strain MAV-2(O)/2) (frag

C:Species: myeloblastosis-associated virus

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 26-Feb-1998

C:Accession: I48613

R:Joliot, V.; Borouhs, K.; Lasserre, F.; Crochet, J.; Dambrine, G.; Smith, R.E.; Per

Virology 195, 812-819, 1993

A:Title: Pathogenic potential of myeloblastosis-associated virus: implication of env

A:Reference number: A48613; MUID:9331743; PMID:8393249

A:Accession: I48613

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-601 <JOL>

C:Superfamily: type C retrovirus env polyprotein

Query Match	5.2%;	Score 7;	DB 2;	Length 601;
Best Local Similarity	100.0%;	Pred. No. 44;		
Matches 7;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 33 ANLTSL 39

DB 454 ANLTSL 460

RESULT 37

VCFPER

env polyprotein - Rous sarcoma virus (strain Prague C)

N:Contains: coat protein gp37; coat protein gp5

C:Species: Rous sarcoma virus

C>Date: 01-Sep-1981 #sequence_revision 17-Dec-1982 #text_change 16-Jun-2000

C:Accession: A03996; S26419; S03602

R:Schwartz, D.; Tizard, R.; Gilbert, W.

Submitted to the Nucleic Acid Sequence Database, September 1982

A:Reference number: A00632

A:Accession: A03996

A:Molecule type: genomic RNA

A:Residues: 1-603 <SCH>

A:Note: as a result of base variations, a different version of this sequence may exist

R:Kashuba, V.I.; Rynditch, A.V.; Dostalova, V.; Hlozaneck, I.; Zubak, S.V.; Kavan, V.

Submitted to the EMBL Data Library, September 1992

A:Description: Molecular cloning and DNA sequence analysis of duck-adapted variant of

A:Reference number: S26417

A:Accession: S26419

A:Molecule type: DNA

A:Residues: 27-133, 'S', '135-145', 'A', '147-157', 'S', '159-312', 'A', '314-324', 'KTT', '328-331', 'S',

A:Cross-references: EMBL:X68524; NID:961903; PID:CAA48536.1; PID:91334938

R:Kashuba, V.I.; Zubak, S.V.; Rynditch, A.V.; Kavan, V.M.; Hlozaneck, I.; Svoboda, J.

Nucleic Acids Res. 17, 2120, 1989
 A:Title: The nucleotide sequence of the region of src gene deletion in transformation-de
 A:Reference number: S03602; MUID:89183615; PMID:2538803
 A:Accession: S03602
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 511-566, 'V', 568-603 <KAW>
 A:Cross-references: EMBL:X13818; NID:961893; PIDN:CAA32051.1; PID:g833169
 C:Genetics:
 A:Gene: env
 C:Superfamily: type C retrovirus env polypeptide
 C:Keywords: polyprotein
 F:65-603/Product: coat protein gp37 #status predicted <P37>
 F:46-603/Product: coat protein gp37 #status predicted <P37>

Query Match
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 ANLTSL 39
 |||||
 Db 456 ANLTSL 462

RESULT 38
 S49969
 Probable membrane protein YIL014w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein YI329.18
 C:Species: Saccharomyces cerevisiae
 C:Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 19-Apr-2002
 C:Accession: S49969
 R:Skellon, J.; Church, C.
 submitted to the EMBL Data Library, December 1994
 A:Reference number: S49951
 A:Accession: S49951
 A:Molecule type: DNA
 A:Residues: 1-630 <SKE>
 A:Cross-references: EMBL:246881; NID:959967; PID:959986; GSPDB:GN00009; MIPS:YIL014w
 C:Genetics:
 A:Gene: SGD:MNT3; MIPS:YIL014w
 A:Cross-references: SGD:S0001276
 A:Map position: 9L
 C:Keywords: transmembrane protein
 F:15-31/Domain: transmembrane #Status predicted <TMM>

Query Match
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 AKISETA 77
 |||||
 Db 280 AKISETA 286

RESULT 39
 T00025
 PSD-95 binding protein - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
 C:Accession: T00025
 R:Kawashima, N.; Takamiya, K.; Sun, J.; Kitabatake, A.; Sobue, K.
 FEBS Lett. 418, 301-304, 1997
 A:Title: Differential expression of isoforms of PSD-95 binding protein (GKAP/SAPAP1) dur
 A:Reference number: 214063; MUID:98089008; PMID:9428732
 A:Accession: T00025
 A:Molecule type: mRNA
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Residues: 1-692 <KAW>
 A:Cross-references: EMBL:AB003594; NID:g2766158; PIDN:BAA24265.1; PID:g2766159

Query Match
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 SITELETT 96
 |||||
 Db 89 SITELETT 95

RESULT 40
 AC1312
 ATP-dependent helicases homolog dling [imported] - listeria monocytogenes (strain EGD-
 C:Species: Listeria monocytogenes
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AC1312
 R:Glaser, P.; Frangenau, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maltournam, A.;
 Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
 A:Title: Comparative genomics of listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AC1312
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-928 <GLA>
 A:Cross-references: GB:NC_003210; PIDN:CAC99977.1; PID:g16411352; GSPDB:GN00177
 C:Genetics:
 A:Experimental source: strain EGD-e
 A:Gene: dling

Query Match
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 IAANKLS 66
 |||||
 Db 520 IAANKLS 526

Search completed: April 28, 2003, 16:15:12
 Job time : 50 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 16:05:17 ; Search time 12 Seconds

(without alignments)
463.152 Million cell updates/sec

Title: US-10-048-196-2

Perfect score: 134

Sequence: 1 MMKILYVATLMTAFTLASG.....SLKPCMETVNAFTVPTTFR 134

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_40:*
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	5.2	157	1	DUP_BRUME
2	7	5.2	174	1	ENV_AVISU
3	7	5.2	246	1	ENV_RSISR
4	7	5.2	257	1	ENV_AVIR3
5	7	5.2	271	1	VDIC_HELPJ
6	7	5.2	281	1	Y149_MYCGE
7	7	5.2	343	1	CATQ_RAT
8	7	5.2	345	1	ABC_HAEIN
9	7	5.2	348	1	SELD_HAEIN
10	7	5.2	352	1	CLTI_MOUSE
11	7	5.2	466	1	DCOR_YEAST
12	7	5.2	486	1	DNAA_BORBU
13	7	5.2	603	1	ENV_RSISR
14	7	5.2	630	1	MNTV_YEAST
15	7	5.2	977	1	DLPI_HUMAN
16	7	5.2	992	1	DLPI_RAT
17	7	5.2	55	1	RU32_PASMU
18	6	4.5	57	1	Y147_UREPA
19	6	4.5	65	1	SECE_SULAC
20	6	4.5	89	1	RSIS_SYN3
21	6	4.5	123	1	MP11_LYMST
22	6	4.5	136	1	YCV4_YEAST
23	6	4.5	137	1	MP12_LYMST
24	6	4.5	138	1	MP15_LYMST
25	6	4.5	144	1	PSB2_ECOLI
26	6	4.5	150	1	RSI3_YEAST
27	6	4.5	159	1	MPAG_ALNGL
28	6	4.5	176	1	DSBG_SHIFL
29	6	4.5	192	1	SC72_YEAST
30	6	4.5	196	1	RETB_CHICK
31	6	4.5	201	1	RETB_HUMAN
32	6	4.5	202	1	YNIC_ECOLI
33	6	4.5	238	1	Y647_HAEIN

34	6	4.5	239	1	YG1P_YEAST
35	6	4.5	246	1	Y564_METTH
36	6	4.5	252	1	YE91_MYCTU
37	6	4.5	254	1	GLTF_ECOLI
38	6	4.5	258	1	RM24_YEAST
39	6	4.5	258	1	YFAP_ECOLI
40	6	4.5	261	1	RF44_HUMAN
41	6	4.5	261	1	YIHW_ECOLI
42	6	4.5	262	1	ZNIB_BICAT
43	6	4.5	269	1	CLB2_CHICK
44	6	4.5	271	1	PD44_HUMAN
45	6	4.5	272	1	CPCE_SYN3
46	6	4.5	272	1	TGIF_HUMAN
47	6	4.5	273	1	ILVE_THEMA
48	6	4.5	273	1	MUR1_AGR5
49	6	4.5	273	1	PSAB_YERPE
50	6	4.5	274	1	PD44_BOVIN
51	6	4.5	274	1	PD44_MOUSE
52	6	4.5	275	1	2253_HUMAN
53	6	4.5	281	1	KD5A_PSEAE
54	6	4.5	284	1	VDIC_HELPJ
55	6	4.5	286	1	YMB1_CAEEL
56	6	4.5	300	1	NC5R_BOVIN
57	6	4.5	308	1	MAUR_KLEPN
58	6	4.5	309	1	RAS1_YEAST
59	6	4.5	311	1	MIAA_HAEIN
60	6	4.5	315	1	T2A1_ANASP
61	6	4.5	316	1	TALB_ECOLI
62	6	4.5	316	1	TALB_SALTY
63	6	4.5	317	1	TAL_YERPE
64	6	4.5	320	1	NOD1_AZOCA
65	6	4.5	329	1	YE29_HELPJ
66	6	4.5	329	1	YE29_HELPJ
67	6	4.5	332	1	G3P_STRAU
68	6	4.5	332	1	Y4PJ_RHSIN
69	6	4.5	334	1	YEPH_ECOLI
70	6	4.5	334	1	CHM1_HUMAN
71	6	4.5	336	1	Y1F6_YEAST
72	6	4.5	337	1	G3P_MYCGE
73	6	4.5	337	1	G3P_MYCPN
74	6	4.5	339	1	ERA_RICCN
75	6	4.5	342	1	G3P_AOUAE
76	6	4.5	345	1	VAOD_YEAST
77	6	4.5	350	1	IL8A_HUMAN
78	6	4.5	357	1	MURC_ANASP
79	6	4.5	360	1	CISP_HEMSP
80	6	4.5	361	1	YCHF_BUCAL
81	6	4.5	367	1	TROD_TREPA
82	6	4.5	369	1	METR_HAEIN
83	6	4.5	374	1	G3PA_CHIRE
84	6	4.5	376	1	YOR3_CAEEL
85	6	4.5	377	1	AROB_AGR5
86	6	4.5	377	1	MYG_SULDI
87	6	4.5	380	1	FOS_MOUSE
88	6	4.5	381	1	FOS_MOUSE
89	6	4.5	385	1	FABI_BRANA
90	6	4.5	387	1	PAB_PEPMA
91	6	4.5	390	1	NEUC_ECOLI
92	6	4.5	410	1	PLGR_MAIZE
93	6	4.5	412	1	MTRC_NEICO
94	6	4.5	414	1	G3PA_CHOOR
95	6	4.5	416	1	G3PA_GRAVE
96	6	4.5	418	1	NTR1_HUMAN
97	6	4.5	419	1	PSG1_HUMAN
98	6	4.5	422	1	NCAP_CHAV
99	6	4.5	424	1	CYAA_STIAU
100	6	4.5	424	1	NTR1_MOUSE
101	6	4.5	424	1	NTR1_MOUSE
102	6	4.5	426	1	HEM1_KANCH
103	6	4.5	428	1	PSG3_HUMAN
104	6	4.5	430	1	PRST_SALTY
105	6	4.5	431	1	DCOR_DATST
106	6	4.5	433	1	AS14_MOUSE

P53223	saccharomyc
O26664	methanobact
P71772	mycobacteri
P28721	eschericchia
P36525	saccharomyc
P76462	eschericchia
Q13156	homo sapien
P32144	eschericchia
P57402	buchnera ap
P07090	gallus gall
Q13162	homo sapien
P73638	synchocyst
Q15583	homo sapien
P74921	thermotoga
O8ue93	agrobacteri
P31523	yesinia pe
O9Bg12	bos taurus
O08807	mus musculu
O75346	homo sapien
O9ZFK4	pseudomonas
O05730	helicobacte
P50093	caenorhabd
P07514	bos taurus
P01119	saccharomyc
P52684	klebsiella
P70803	haemophilus
P44495	haemophilus
P30148	eschericchia
Q8xg45	salmonella
O8xln2	yesinia pe
O07756	azorhizobiu
O9Z115	helicobacte
O25971	helicobacte
O59600	streptomyce
P55664	rhizobium s
P39836	eschericchia
O75829	homo sapien
P40493	saccharomyc
P47543	mycoplasma
P73538	mycoplasma
O921a9	rickettsia
O67161	aquifex aeo
P32366	saccharomyc
P25024	homo sapien
O8Z123	anabaena sp
P43156	hemerocall
P57288	buchnera ap
P96419	treponema p
P44502	haemophilus
P50362	chlamydomon
O09315	caenorhabd
Q8u9v0	agrobacteri
Q01966	suliculus di
P01101	mus musculu
P01102	fbj murine
P80030	brassica na
O51911	peptostrept
Q47400	eschericchia
P26216	zea mays (m
P43505	neisseria g
P34919	chondrus cr
P30724	gracilaria
P30989	homo sapien
P11464	homo sapien
P40137	stigmatalia
P40137	stigmatalia
O88319	mus musculu
P20789	retus norv
P42808	xanthomonas
Q16537	homo sapien
O40193	salmonella
P50134	datura stra
Q8vhs7	mus musculu

107	6	4.5	434	1	YAOA_SCHPO	Q10089 schizosacch	180	6	4.5	734	1	UN36_CAEEL	P34374 caenorhabd
108	6	4.5	436	1	VE2_HP22	P50768 human papil	181	6	4.5	739	1	PURL_LACLA	O9cfe8 lactococcus
109	6	4.5	439	1	NTT4_HUMAN	O9h1v8 homo sapien	182	6	4.5	739	1	PURL_LACLA	O9ab06 lactococcus
110	6	4.5	440	1	TYPH_ECOLI	P07650 escherichia	183	6	4.5	767	1	AMVH_SACDI	O4065 saccharomyc
111	6	4.5	445	1	TIG_RICCN	Q29200 rickettsia	184	6	4.5	768	1	AMVH_SACDI	P22760 saccharomyc
112	6	4.5	451	1	HAK_SCHMA	Q26609 schistosoma	185	6	4.5	774	1	SFE_LAMBD	P03764 bacterioph
113	6	4.5	461	1	TRPE_THEMEA	O08653 thermotoga	186	6	4.5	784	1	Y063_CAEEL	P34603 caenorhabd
114	6	4.5	481	1	ALGA_PSEAE	P07874 p alginatae	187	6	4.5	828	1	SOX6_HUMAN	P35712 homo sapien
115	6	4.5	513	1	YOGP_BACPU	P18267 bacillus su	188	6	4.5	830	1	NAPA_YEAST	P21268 saccharomyc
116	6	4.5	513	1	TRPE_BACPU	Q1eth5 methanosarc	189	6	4.5	831	1	NAPA_RHOSH	O51176 rhodobacter
117	6	4.5	515	1	LEI2_METNA	P53968 saccharomyc	190	6	4.5	840	1	PHL1_HUMAN	P80108 homo sapien
118	6	4.5	516	1	KTR5_YEAST	P31603 rous sarcom	191	6	4.5	850	1	DEXT_STRKU	Q54443 streptococc
119	6	4.5	522	1	KTR5_YEAST	O15742 homo sapien	192	6	4.5	857	1	WGLB_EBV	Q93c8 pyrococcus
120	6	4.5	523	1	NAB2_HUMAN	P16052 rous sarcom	193	6	4.5	880	1	RA50_PVRAB	P43407 caenorhabd
121	6	4.5	525	1	NAB2_HUMAN	O61177 mus musculu	194	6	4.5	886	1	YKRI_CAEEL	O14940 homo sapien
122	6	4.5	525	1	NAB2_HUMAN	Q59200 corynebacte	195	6	4.5	896	1	NAH5_HUMAN	P26420 lactobacill
123	6	4.5	526	1	ASPA_CORGL	P00525 avian sarco	196	6	4.5	901	1	SVY_LACRA	Q94f21 aeropyrum p
124	6	4.5	526	1	SRC_AVISR	P25020 rous sarcom	197	6	4.5	919	1	RA50_LAEPE	P80012 bos taurus
125	6	4.5	526	1	SRC_RSVA1	P00526 rous sarcom	198	6	4.5	937	1	VWF_BOVIN	Q93c8 pyrococcus
126	6	4.5	526	1	SRC_RSVA1	P00524 rous sarcom	199	6	4.5	949	1	STB4_YEAST	P43012 epstein-bar
127	6	4.5	526	1	SRC_RSVA1	P00524 rous sarcom	200	6	4.5	953	1	SVY_VIBCH	Q93c8 pyrococcus
128	6	4.5	528	1	UDBJ_MACFA	O9xt55 macaca fasc	201	6	4.5	954	1	SVY_VIBCH	Q93c8 pyrococcus
129	6	4.5	529	1	VGLE_SVS	P04849 similan viru	202	6	4.5	956	1	SVY_TREPA	P80012 bos taurus
130	6	4.5	531	1	SRC1_XENLA	P13115 xenopus lae	203	6	4.5	971	1	US51_MOUSE	O83988 haemophilus
131	6	4.5	531	1	TYD2_PAPSO	P13116 xenopus lae	204	6	4.5	972	1	US51_MOUSE	O08810 mus musculu
132	6	4.5	532	1	SRC_CHICK	P54769 papaver som	205	6	4.5	980	1	BOB1_YEAST	O15029 homo sapien
133	6	4.5	535	1	SRC_HUMAN	P12931 homo sapien	206	6	4.5	990	1	ENV_OMVVS	P38041 saccharomyc
134	6	4.5	535	1	SRC_HUMAN	Q9wud9 rattus norv	207	6	4.5	993	1	SM14_SCHPO	P16899 ovine lentl
135	6	4.5	538	1	VGLE_NDMP1	P19716 mumps virus	208	6	4.5	994	1	MYSA_DICTDI	P22467 dictyostell
136	6	4.5	538	1	VGLE_NDMP1	P19716 mumps virus	209	6	4.5	1002	1	MYSA_DICTDI	P22467 dictyostell
137	6	4.5	538	1	VGLE_NDMP1	P19716 mumps virus	210	6	4.5	1002	1	YAY3_ANASP	O50439 mycobacteri
138	6	4.5	538	1	VGLE_NDMP1	P19716 mumps virus	211	6	4.5	1002	1	YAY3_ANASP	O50439 mycobacteri
139	6	4.5	540	1	VGLE_NDMP1	P19716 mumps virus	212	6	4.5	1008	1	DPOL_HSV2	P58612 anabaena sp
140	6	4.5	540	1	VGLE_NDMP1	P19716 mumps virus	213	6	4.5	1076	1	MLLA_MYCLE	O49619 equine herp
141	6	4.5	551	1	VGLE_P12HG	P03481 mumps virus	214	6	4.5	1083	1	RPOB_ASTLO	P27059 astasia lon
142	6	4.5	551	1	VGLE_P12HG	P03481 mumps virus	215	6	4.5	1123	1	V120_HSV11	O00268 homo sapien
143	6	4.5	551	1	VGLE_P12HG	P03481 mumps virus	216	6	4.5	1172	1	CU14_SCHPO	P10221 herpes simp
144	6	4.5	553	1	VGLE_P12HT	P27286 human paral	217	6	4.5	1205	1	YLJ9_CAEEL	P41003 schizosacch
145	6	4.5	553	1	VGLE_P12HT	P27286 human paral	218	6	4.5	1264	1	MOUR_ECOLI	P34372 caenorhabd
146	6	4.5	553	1	VGLE_P12HT	P27286 human paral	219	6	4.5	1276	1	BXD_CLOBO	P33345 escherichia
147	6	4.5	553	1	VGLE_P12HT	P27286 human paral	220	6	4.5	1335	1	XDH_DROME	P10321 clostridium
148	6	4.5	553	1	VGLE_P12HT	P27286 human paral	221	6	4.5	1337	1	DEXT_STRODO	P87365 streptococc
149	6	4.5	553	1	VGLE_P12HT	P27286 human paral	222	6	4.5	1337	1	DEXT_STRODO	P87365 streptococc
150	6	4.5	553	1	VGLE_P12HT	P27286 human paral	223	6	4.5	1342	1	YDMS_SCHPO	P22811 drosophila
151	6	4.5	553	1	VGLE_P12HT	P27286 human paral	224	6	4.5	1343	1	XDH_DROPS	P17482 drosophila
152	6	4.5	553	1	VGLE_P12HT	P27286 human paral	225	6	4.5	1344	1	TOP2_BOMMO	P91711 drosophila
153	6	4.5	553	1	VGLE_P12HT	P27286 human paral	226	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
154	6	4.5	553	1	VGLE_P12HT	P27286 human paral	227	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
155	6	4.5	553	1	VGLE_P12HT	P27286 human paral	228	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
156	6	4.5	553	1	VGLE_P12HT	P27286 human paral	229	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
157	6	4.5	553	1	VGLE_P12HT	P27286 human paral	230	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
158	6	4.5	553	1	VGLE_P12HT	P27286 human paral	231	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
159	6	4.5	553	1	VGLE_P12HT	P27286 human paral	232	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
160	6	4.5	553	1	VGLE_P12HT	P27286 human paral	233	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
161	6	4.5	553	1	VGLE_P12HT	P27286 human paral	234	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
162	6	4.5	553	1	VGLE_P12HT	P27286 human paral	235	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
163	6	4.5	553	1	VGLE_P12HT	P27286 human paral	236	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
164	6	4.5	553	1	VGLE_P12HT	P27286 human paral	237	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
165	6	4.5	553	1	VGLE_P12HT	P27286 human paral	238	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
166	6	4.5	553	1	VGLE_P12HT	P27286 human paral	239	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
167	6	4.5	553	1	VGLE_P12HT	P27286 human paral	240	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
168	6	4.5	553	1	VGLE_P12HT	P27286 human paral	241	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
169	6	4.5	553	1	VGLE_P12HT	P27286 human paral	242	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
170	6	4.5	553	1	VGLE_P12HT	P27286 human paral	243	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
171	6	4.5	553	1	VGLE_P12HT	P27286 human paral	244	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
172	6	4.5	553	1	VGLE_P12HT	P27286 human paral	245	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
173	6	4.5	553	1	VGLE_P12HT	P27286 human paral	246	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
174	6	4.5	553	1	VGLE_P12HT	P27286 human paral	247	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
175	6	4.5	553	1	VGLE_P12HT	P27286 human paral	248	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
176	6	4.5	553	1	VGLE_P12HT	P27286 human paral	249	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
177	6	4.5	553	1	VGLE_P12HT	P27286 human paral	250	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
178	6	4.5	553	1	VGLE_P12HT	P27286 human paral	251	6	4.5	1344	1	XDH_DROPS	P91711 drosophila
179	6	4.5	553	1	VGLE_P12HT	P27286 human paral	252	6	4.5	1344	1	XDH_DROPS	P91711 drosophila

253	5	3.7	65	1	MYX1_CRODU	P24331	crocalus du	326	5	3.7	117	1	Y117_HAEDU	O30825	haemophilus
254	5	3.7	65	1	MYX3_CRODU	P24433	crocalus du	327	5	3.7	117	1	YCX2_CHIRE	P05722	chlamydomon
255	5	3.7	66	1	ITR2_ASCSU	P01049	ascaris suu	328	5	3.7	119	1	RNPA_HAETN	P43030	haemophilus
256	5	3.7	66	1	RL29_BACST	P04457	baclillus st	329	5	3.7	119	1	RNPA_PASMU	P57915	pasteurella
257	5	3.7	66	1	RL29_THEMA	P39814	thermotoga	330	5	3.7	119	1	TYAE_BACSU	O32227	baclillus su
258	5	3.7	66	1	YSC6_YEREN	Q01246	yersinia en	331	5	3.7	120	1	NLR6_DAUCA	P27631	daucus caro
259	5	3.7	69	1	RK32_PINTH	P41651	pinus thunb	332	5	3.7	120	1	YF65_ESCHIR	P76515	escherichia
260	5	3.7	74	1	NINY_BP222	O38668	bacterioph	333	5	3.7	120	1	YJ08_YEAST	P47106	saccharomyc
261	5	3.7	75	1	UL91_HSVSA	Q01010	herpesvirus	334	5	3.7	121	1	H1A_PLADU	P06694	platyneris
262	5	3.7	76	1	EC_MAIZE	P43401	zea mays (m	335	5	3.7	121	1	RS13_TREPA	O83240	treponema p
263	5	3.7	78	1	YCX1_DICDH	P30160	dictyola di	336	5	3.7	122	1	RBR4_THETN	O81838	thermoanaer
264	5	3.7	79	1	EXY3_LACLA	O9c8h3	lactococcus	337	5	3.7	122	1	YCCV_ECOLI	P75817	escherichia
265	5	3.7	80	1	EXY5_PSEAE	O9hwy5	pseudomonas	338	5	3.7	123	1	C59A_MOUSE	O55186	mus musculu
266	5	3.7	81	1	YEE9_SCHPO	O13825	schizosacch	339	5	3.7	123	1	NBBM_GAEEL	P90789	caenorhabdi
267	5	3.7	82	1	YBDJ_ECOLI	P77506	escherichia	340	5	3.7	123	1	PHF5_DESVH	P07603	desulfovibr
268	5	3.7	82	1	YCX1_VICFA	P08889	vicia faba	341	5	3.7	123	1	YJ54_YEAST	P46884	saccharomyc
269	5	3.7	87	1	Z154_HUMAN	O9c4d3	mycobacteri	342	5	3.7	124	1	YOB4_ECOLI	P76729	escherichia
270	5	3.7	87	1	Z154_HUMAN	O13106	homo sapien	343	5	3.7	125	1	CP03_HUMAN	O95177	homo sapien
271	5	3.7	88	1	MX31_YEAST	P23059	saccharomyc	344	5	3.7	125	1	RS13_RICCN	O92498	schistosoma
272	5	3.7	89	1	Y12E_BPT4	P33084	bacterioph	345	5	3.7	125	1	RS13_RICPR	O92c57	rickettsia
273	5	3.7	89	1	Y565_SULTO	O975e9	sulfolobus	346	5	3.7	126	1	DESR_DESBR	O46495	desulfifarcu
274	5	3.7	90	1	DBH_RHIME	P02344	rhizobium m	347	5	3.7	126	1	SECG_PSEST	P95577	pseudomonas
275	5	3.7	90	1	VER7_HPV29	P50784	human papil	348	5	3.7	126	1	SYN2_MOUSE	O64332	mus musculu
276	5	3.7	91	1	C554_THINE	P25338	thiobacilli	349	5	3.7	127	1	RL20_STRCO	O88058	streptomyce
277	5	3.7	91	1	SY05_CAVPO	P97272	cavia porce	350	5	3.7	127	1	SECE_ECOLI	P16920	escherichia
278	5	3.7	92	1	SY01_MOUSE	P10146	mus musculu	351	5	3.7	127	1	SECE_SALTY	O919k1	salmonella
279	5	3.7	93	1	NEU1_ANSAN	P35519	anser anser	352	5	3.7	127	1	YHHH_ECOLI	P28811	escherichia
280	5	3.7	93	1	Y974_TREPA	O83939	treponema p	353	5	3.7	128	1	AZ0R_PSEDE	P00283	pseudomonas
281	5	3.7	94	1	DBH_HELPJ	O92108	helicobacte	354	5	3.7	129	1	AAK2_PIG	O28948	sus scrofa
282	5	3.7	94	1	DBH_HELPJ	O25506	helicobacte	355	5	3.7	129	1	Y389_MYCPN	P46793	dictyostell
283	5	3.7	95	1	GLRX_RICPR	O92dW1	rickettsia	356	5	3.7	129	1	Y389_MYCPN	P75508	mycoplasma
284	5	3.7	96	1	E111_ADEMI	P12533	mouse adeno	357	5	3.7	131	1	YOHF_BACSU	P54514	baclillus su
285	5	3.7	97	1	SPAC_BPT4	P39230	bacterioph	358	5	3.7	132	1	F802_SCHMA	P16464	schistosoma
286	5	3.7	97	1	YXRB_CALSA	P23554	caldocellum	359	5	3.7	132	1	PFDA_PYRAE	O8zt2f	pyrobaculum
287	5	3.7	98	1	C50C_THINE	P45888	thiobacilli	360	5	3.7	132	1	YD69_SCHPO	Q10321	schizosacch
288	5	3.7	98	1	NULM_BOVIN	P03902	bos taurus	361	5	3.7	133	1	CRG1_RANTE	P02530	rana tempor
289	5	3.7	98	1	NULM_BOVIN	O03203	ceratotheri	362	5	3.7	133	1	RISH_HALNI	O9hms5	halobacteri
290	5	3.7	98	1	NULM_EQUAS	P92483	equus asinu	363	5	3.7	133	1	VEGH_OREN2	P52884	orf virus (
291	5	3.7	98	1	NULM_GORGO	O34572	gorilla gor	364	5	3.7	134	1	ASER_BOVIN	P29392	bos taurus
292	5	3.7	98	1	NULM_HIPAM	O92zy3	hipopotamu	365	5	3.7	134	1	RL32_DROAC	O94460	drosofila
293	5	3.7	98	1	NULM_HORSE	P46858	equus cabal	366	5	3.7	134	1	RL32_HUMAN	P02433	homo sapien
294	5	3.7	98	1	NULM_HUMAN	P03901	homo sapien	367	5	3.7	135	1	CASK_EOUGR	O28400	equus grevy
295	5	3.7	98	1	NULM_LATCH	O03172	latimeria c	368	5	3.7	135	1	NOLV_RHIFR	P33211	rhizobium f
296	5	3.7	98	1	NULM_PANTR	O37809	pan troglod	369	5	3.7	135	1	HIS3_METVA	O50837	methanococc
297	5	3.7	98	1	NULM_PHOVI	O00544	phoca vitul	370	5	3.7	136	1	SR14_HUMAN	P37108	homo sapien
298	5	3.7	98	1	NULM_SHEEP	O78154	ovis aries	371	5	3.7	137	1	NDK_CAUJE	O919t7	campylobact
299	5	3.7	99	1	SY08_PIG	P49873	sus scrofa	372	5	3.7	137	1	NIRB_KLEOX	P56765	klebsiella
300	5	3.7	100	1	BOP1_BOVIN	P83107	bos taurus	373	5	3.7	137	1	THNS_HORVU	P09617	hordium vul
301	5	3.7	100	1	URE3_YEREN	P31496	yersinia en	374	5	3.7	137	1	VI36_FOMPV	O85811	foxiapox vit
302	5	3.7	100	1	URE3_YERPE	P52315	yersinia pe	375	5	3.7	138	1	RM16_CHOCR	P48955	chonduris cr
303	5	3.7	101	1	TAT_HY1U4	P24738	human timan	376	5	3.7	139	1	ATPD_NEUCR	P56525	neurospora
304	5	3.7	102	1	RS10_THETN	O817v3	thermoanaer	377	5	3.7	139	1	IL5_MACEU	O9x591	macrospora eu
305	5	3.7	102	1	YVCB_BACSU	P33600	baclillus su	378	5	3.7	139	1	RT12_MOUSE	O35680	mus musculu
306	5	3.7	104	1	RS10_THETH	P80375	thermus the	379	5	3.7	139	1	YN06_YEAST	P55842	saccharomyc
307	5	3.7	105	1	YNC8_YEAST	P53967	saccharomyc	380	5	3.7	140	1	CV3B_DESAR	P46691	desulfovibr
308	5	3.7	106	1	RL11_HAEDU	O32613	haemophilus	381	5	3.7	141	1	HBA1_TACAC	P01977	tachygyossu
309	5	3.7	106	1	YAVC_RHISN	O53211	rhizobium s	382	5	3.7	141	1	HBA1_CHRPI	P07423	ursus marit
310	5	3.7	106	1	YMI0_CAREL	P34485	caenorhabdi	383	5	3.7	141	1	HBA_URSPA	P58836	methanopyru
311	5	3.7	109	1	PER_RHACO	O26231	rihaoleitis	384	5	3.7	141	1	HIS3_METKA	O9p1p1	chlamydia m
312	5	3.7	110	1	SR14_CANFA	P16255	canis famli	385	5	3.7	141	1	NDK_GILMU	O84508	chlamydia t
313	5	3.7	110	1	SR14_MOUSE	P16254	mus musculu	386	5	3.7	141	1	RE67_PYRAB	O9v159	pyrococcus
314	5	3.7	111	1	Y546_SYNY3	O55397	synecocyst	387	5	3.7	141	1	RE67_PYRAB	P42180	pyrococcus
315	5	3.7	112	1	MTBB_METKA	O32866	methanopyru	388	5	3.7	141	1	RE67_PYRAB	O59256	pyrococcus
316	5	3.7	113	1	KV2F_MOUSE	P01630	mus musculu	389	5	3.7	141	1	RE67_PYRAB	P44331	haemophilus
317	5	3.7	114	1	PRD2_SCHPO	P12733	schizosacch	390	5	3.7	143	1	RL11_HAETN	AT92_BOVIN	AT92_SHEEP
318	5	3.7	115	1	UL96_HCMTA	O9utcs	haloarcula	391	5	3.7	143	1	AT92_BOVIN	P07926	bos taurus
319	5	3.7	115	1	UL96_HCMTA	P16787	human cytom	392	5	3.7	144	1	NDK_CHLUP	O06056	ovis aries
320	5	3.7	116	1	YOB9_STRCO	O9xa13	streptomyce	393	5	3.7	144	1	NDK_CHLUP	O97175	chlamydia p
321	5	3.7	116	1	NLRP_GERRY	O39794	gerbera hyb	394	5	3.7	144	1	YPHP_BACSU	P54170	baclillus su
322	5	3.7	116	1	RL17_HELPY	P56042	helicobacte	395	5	3.7	146	1	LYCV_BPAPS	O91c15	bacterioph
323	5	3.7	116	1	WNSA_CHDEA	P28109	chen caerul	396	5	3.7	146	1	LPDD_ECOLI	P36647	escherichia
324	5	3.7	116	1	WNSA_EUMSK	P28117	eumeces ski	397	5	3.7	146	1	RL15_METHH	O26133	methanobact
325	5	3.7	116	1	WNSA_MEIGA	P28128	meleagris g	398	5	3.7	146	1	YK81_MYCTU	O10689	mycobacteri

545	5	3.7	200	1	SODM_GANMT	092429	ganoderma m	618	5	3.7	217	1	HRCR_RALSO	052488	ralstonia s
546	5	3.7	201	1	NING_BPH33	0924x1	bacterioph	619	5	3.7	217	1	IF3_STRCO	088600	streptomyce
547	5	3.7	201	1	NING_BPH19	048427	bacterioph	620	5	3.7	217	1	RS3_BACST	088609	bacillus st
548	5	3.7	201	1	RETR_PIG	P27485	sus scrofa	621	5	3.7	218	1	Y336_STNY3	P23056	synecocyst
549	5	3.7	201	1	YCD1_HUMAN	09y3d2	homo sapien	622	5	3.7	218	1	YV56_CAEEL	022019	caenorhabd1
550	5	3.7	202	1	IF3_MYCLE	09cc22	mycobacteri	623	5	3.7	219	1	EGU5_CAEEL	P17486	caenorhabd1
551	5	3.7	202	1	LIF_BOVIN	027956	bos taurus	624	5	3.7	219	1	GL22_ARATH	09f427	arabidopsis
552	5	3.7	202	1	R13A_HUMAN	P40429	homo sapien	625	5	3.7	219	1	MGH1_HUMAN	09h313	homo sapien
553	5	3.7	202	1	R13A_MOUSE	P19253	mus musculu	626	5	3.7	219	1	NOTA_BRAEL	P50329	bradyrhizob
554	5	3.7	202	1	R13A_MOUSE	P35427	rattus norv	627	5	3.7	220	1	SODE_SOLCS	004997	solidago ca
555	5	3.7	202	1	RH01_SCHPO	009914	schizosacch	628	5	3.7	220	1	UPAS_RAT	P51573	rattus norv
556	5	3.7	202	1	WRB2_RHIME	093012	rhizobium m	629	5	3.7	220	1	YA78_SCHPO	009761	schizosacch
557	5	3.7	203	1	CHPE_PSEAE	087005	pseudomonas	630	5	3.7	220	1	YF09_HELPJ	092161	helicobacte
558	5	3.7	203	1	IDI_MYCTU	P72002	mycobacteri	631	5	3.7	220	1	YF09_HELPJ	026309	helicobacte
559	5	3.7	203	1	NING_BPP22	038667	bacterioph	632	5	3.7	221	1	BRF1_EBV	P03328	epstein-bar
560	5	3.7	203	1	YDAM_ECOLI	P76066	escherichia	633	5	3.7	221	1	TERF_PHOSU	P97516	phodopus su
561	5	3.7	204	1	NING_BPP21	09x194	bacterioph	634	5	3.7	221	1	YSC1_YERPE	000928	yersinia pe
562	5	3.7	204	1	NING_LAMB	P03770	bacterioph	635	5	3.7	221	1	RNC_ZYMO	0925u2	zymomonas m
563	5	3.7	204	1	PS21_ARATH	023714	arabidopsis	636	5	3.7	222	1	RNS3_PYPY	0925u2	zymomonas m
564	5	3.7	204	1	SRY_GORGO	P48046	gorilla gor	637	5	3.7	222	1	Y815_RICPR	092ce2	ricetstia
565	5	3.7	204	1	SRY_GORGO	005066	homo sapien	638	5	3.7	222	1	GLIA_ARATH	09ma9	arabidopsis
566	5	3.7	204	1	SRY_HUMAN	028447	hylobates l	639	5	3.7	223	1	RL4_MYCBO	064045	mycobacteri
567	5	3.7	204	1	XPAC_BACSU	P37467	bacillus su	640	5	3.7	223	1	YCS2_BPM2	064226	mycobacteri
568	5	3.7	205	1	CYOC_BUCAL	P57542	buchnera ap	641	5	3.7	223	1	YSC1_YEREN	001253	yersinia en
569	5	3.7	205	1	UBA1_CANAL	P52495	candida alb	642	5	3.7	224	1	PUR0_MYCLE	005756	mycobacteri
570	5	3.7	205	1	Y4JH_RHISN	P55508	rhizobium s	643	5	3.7	224	1	Y209_LISMO	P33381	listeria mo
571	5	3.7	206	1	NOO5_PARDE	P29917	paracoccus	644	5	3.7	224	1	YCS5_ECOLI	P33792	escherichia
572	5	3.7	206	1	RNMG_ECO57	P58445	escherichia	645	5	3.7	224	1	YCS5_ECOLI	P33792	escherichia
573	5	3.7	206	1	RNMG_ECOLI	P77285	escherichia	646	5	3.7	224	1	YN89_YEAST	P53721	saccharomyc
574	5	3.7	207	1	CLPP_ECOLI	P19245	escherichia	647	5	3.7	225	1	YN99_ARCFU	092d72	archaeoglob
575	5	3.7	207	1	CLPP_SALTY	091c07	salmomella	648	5	3.7	225	1	HX47_MORSA	09pd44	morone saxa
576	5	3.7	207	1	GIDB_NEIMA	091x38	neisseria m	649	5	3.7	225	1	PUR0_CORAM	09trx0	corynebacte
577	5	3.7	207	1	RR4_CHLYU	09k1q3	neisseria m	650	5	3.7	226	1	VP4_BPHK7	P49860	bacterioph
578	5	3.7	207	1	SNAT_MESNU	P32975	chlorella v	651	5	3.7	226	1	ATP6_PONPP	P92719	pongo pygma
579	5	3.7	207	1	TER8_PASMD	P51561	pasteurella	652	5	3.7	227	1	RNS7_PYPY	080325	pyrus pyrif
580	5	3.7	208	1	GTP_CAEEL	P10299	caenorhabd1	653	5	3.7	227	1	CBBY_RHOCA	083513	rhodobacter
581	5	3.7	208	1	KTHY_AERPE	09ya48	aeropyrum p	654	5	3.7	227	1	CIS3_YEAST	P47001	saccharomyc
582	5	3.7	208	1	NOLV_RHISN	P55716	rhizobium s	655	5	3.7	227	1	COX2_CHEME	P98027	cheltogaleu
583	5	3.7	208	1	NOLV_RHISN	P36393	sus scrofa	656	5	3.7	227	1	COX2_DADMA	P98027	cheltogaleu
584	5	3.7	209	1	RL4_MYCGA	052333	mycoplasma	657	5	3.7	227	1	COX2_GALSE	P98032	daubentonia
585	5	3.7	210	1	COX2_TRYCR	P98023	typanosoma	658	5	3.7	227	1	COX2_NYCCO	P98039	nycticebus
586	5	3.7	210	1	KCY_AGR5	08u819	agrobacteri	659	5	3.7	227	1	COX2_PROPA	P98042	propithecus
587	5	3.7	210	1	RL13_BRARE	090210	brachydantio	660	5	3.7	227	1	GLI32_ARATH	037649	rousetus l
588	5	3.7	210	1	RL13_CRIGR	092313	cricetulus	661	5	3.7	227	1	PR1_RABIT	092372	arabidopsis
589	5	3.7	210	1	RL13_HUMAN	P26373	homo sapien	662	5	3.7	227	1	PSAF_CHLRE	028632	oryctolagus
590	5	3.7	210	1	RL13_HUMAN	090375	ictalurus p	663	5	3.7	227	1	RNS5_PYPY	P12356	chlamydomon
591	5	3.7	210	1	RL13_HUMAN	P47963	mus musculu	664	5	3.7	227	1	SRY_CALTA	P93460	pyrus pyrif
592	5	3.7	210	1	RL13_MOUSE	P41123	rattus norv	665	5	3.7	228	1	ATP6_PELSU	P51501	callithrix
593	5	3.7	210	1	RS5_PODCA	008364	podocoryne	666	5	3.7	228	1	TRPC_PYROO	076675	pyrococcus
594	5	3.7	211	1	CLD7_MOUSE	092421	mus musculu	667	5	3.7	228	1	TRPD_PLAFA	P32086	plasmiodium
595	5	3.7	211	1	VRAC_HUMAN	013103	homo sapien	668	5	3.7	229	1	SRY_BISBO	003225	bos taurus
596	5	3.7	211	1	VRAC_HUMAN	003545	bacterioph	669	5	3.7	229	1	SRY_BOVIN	027949	bison bonas
597	5	3.7	211	1	YCY5_ASTLO	P14758	astasia lon	670	5	3.7	230	1	RUM1_SCHPO	P04380	schizosacch
598	5	3.7	211	1	YCY5_ASTLO	P75410	mycoplasma	671	5	3.7	230	1	TOLO_ECOLI	P05828	schizosacch
599	5	3.7	212	1	KCY_RHIME	092sv4	rhizobium m	672	5	3.7	231	1	YB06_YEAST	P38256	saccharomyc
600	5	3.7	212	1	KCY_RHIME	092sv4	rhizobium m	673	5	3.7	231	1	NEP1_CAEEL	09xx15	caenorhabd1
601	5	3.7	213	1	GIDB_RHIME	092k43	rhizobium m	674	5	3.7	231	1	YR01_CAEEL	010014	caenorhabd1
602	5	3.7	213	1	RECU_LACIA	09c122	lactococcus	675	5	3.7	232	1	DNAD_BACSU	P39787	baclillus su
603	5	3.7	213	1	URK_ECOLI	P31218	escherichia	676	5	3.7	232	1	HB23_MOUSE	P04231	mus musculu
604	5	3.7	213	1	URK_ECOLI	08xey2	salmonella	677	5	3.7	232	1	YB06_YEAST	006816	stigmatella
605	5	3.7	213	1	URK_YERPE	08xey2	versinia pe	678	5	3.7	232	1	YB06_YEAST	P01013	galius galli
606	5	3.7	213	1	YOUN_CAEEL	P34658	caenorhabd1	679	5	3.7	232	1	YB06_YEAST	P30616	galius galli
607	5	3.7	214	1	UBL_APLCA	001391	aplysia cal	680	5	3.7	233	1	SRY_PONPP	P36566	saccharomyc
608	5	3.7	214	1	WIF_SIVS4	P12505	simian limu	681	5	3.7	233	1	TP1_TREPA	028783	pongo pygma
609	5	3.7	214	1	WIF_SIVS4	P19506	simian limu	682	5	3.7	234	1	TRAR_AGRV1	P54294	treponema p
610	5	3.7	214	1	YJ82_AQDAE	067790	aquifex aeo	683	5	3.7	234	1	TRAR_AGRV1	P33909	agrobacteri
611	5	3.7	215	1	THIE_LACIA	09c948	lactococcus	684	5	3.7	234	1	TRMD_RICPR	092e37	ricetstia
612	5	3.7	215	1	YC66_RICCN	092957	ricketstia	685	5	3.7	234	1	YAG5_STALU	P55178	staphylococ
613	5	3.7	215	1	YHID_ECOLI	P26606	escherichia	686	5	3.7	234	1	YD58_SCHPO	056230	schizosacch
614	5	3.7	216	1	HRAL_HARIN	003846	haemophilus	687	5	3.7	235	1	RS4E_THEAC	09bw44	thermoplasm
615	5	3.7	216	1	KCY_RHILLO	098cc2	rhizobium l	688	5	3.7	235	1	RS4E_THEAC	09bw44	thermoplasm
616	5	3.7	216	1	YBFG_ECOLI	P37003	escherichia	689	5	3.7	236	1	BCL2_CRITO	093j48	cricetulus
617	5	3.7	216	1	YPT6_CHLRE	039572	chlamydomon	690	5	3.7	236	1	BCL2_MOUSE	P10417	mus musculu


```

983 5 3.7 303 1 MURB_BACSU P18579 bacillus su
984 5 3.7 303 1 PANE_HALN1 O9hr10 halobacteri
985 5 3.7 303 1 PUR7_PICAN O96v66 pichia angu
986 5 3.7 305 1 ERA_VIRCH O9kpx3 vibrio chol
987 5 3.7 305 1 HEM3_XYLEA O9pcx7 xylella fas
988 5 3.7 305 1 SAX1_MOUSE P42580 mus musculu
989 5 3.7 305 1 TOU4_CAEEL P34627 caenorhabdi
990 5 3.7 306 1 KHSE_SULTO O97547 sulfobolus
991 5 3.7 306 1 Y910_TREPA O83880 treponema p
992 5 3.7 306 1 YLIC_ECOLI O83880 escherichia
993 5 3.7 307 1 BLAC_MYCTU O10670 mycobacteri
994 5 3.7 307 1 ISPH_CHITR O84867 chlamydia t
995 5 3.7 307 1 KHSE_DETRA O9trus delinococcus
996 5 3.7 307 1 PM34_HUMAN O43808 homo sapien
997 5 3.7 307 1 PM34_MOUSE O70579 mus musculu
998 5 3.7 307 1 RBH3_CAEEL O20365 caenorhabdi
999 5 3.7 307 1 RNHL_YEAST P53942 saccharomyc
1000 5 3.7 307 1 TRUB_PASMU O9cmq7 pasteurella

```

ALIGNMENTS

RESULT 1

DUT_BRUME STANDARD; PRT; 157 AA.

AC 08YT14;

15-JUN-2002 (Rel. 41, Created)

15-JUN-2002 (Rel. 41, Last sequence update)

15-JUN-2002 (Rel. 41, Last annotation update)

Deoxyuridine 5'-triphosphate nucleotidohydrolase (EC 3.6.1.23)

(dUTPase) (dUTP pyrophosphatase).

DUT OR BME10358.

Brucella melitensis.

Bacteria: Proteobacteria: alpha subdivision: Rhizobiaceae group;

Brucellaceae: Brucella.

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

NCBI_TaxID=29459;

Query Match 5.2%; Score 7; DB 1; Length 157;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 AKISETA 77
Db 139 AKISETA 145

RESULT 2

ENV_AVISU STANDARD; PRT; 174 AA.

AC P03398;

21-JUL-1986 (Rel. 01, Created)

21-JUL-1986 (Rel. 01, Last sequence update)

01-APR-1990 (Rel. 14, Last annotation update)

Coat protein gp37 (ENV protein gp37).

ENV.

OS Avian sarcoma virus (strain UR2).

OC Viruses; Retroviridae; Retroviridae; Avian type C retroviruses.

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

NCBI_TaxID=11883;

Query Match 5.2%; Score 7; DB 1; Length 174;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 ANLTSL 39
Db 21 ANLTSL 27

RESULT 3

ENV_RSVR STANDARD; PRT; 246 AA.

AC P03397;

21-JUL-1986 (Rel. 01, Created)

21-JUL-1986 (Rel. 01, Last sequence update)

15-JUL-1999 (Rel. 38, Last annotation update)

Membrane protein gp37 (Fragment).

ENV.

OS Rous sarcoma virus (strain Schmidt-Ruppin).

OC Viruses; Retroviridae; Retroviridae; Alpharetrovirus.

NCBI_TaxID=11889;

NCBI_TaxID=11889;

NCBI_TaxID=11889;

NCBI_TaxID=11889;

NCBI_TaxID=11889;

NCBI_TaxID=11889;

NCBI_TaxID=11889;

NCBI_TaxID=11889;

NCBI_TaxID=11889;

NCBI_TaxID=11889;

RP SEQUENCE FROM N.A.
 RX MEDLINE=83059858; PubMed-6292477;
 RA Takeya T., Feldman R.A., Hanafusa H.;
 RT "DNA sequence of the viral and cellular src gene of chickens. 1.
 RT Complete nucleotide sequence of an EcoRI fragment of recovered avian
 RT sarcoma virus which codes for gp37 and pp60src";
 RL J. Virol. 44:1-11(1982).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: V01169; CAA24494.1; -
 DR EMBL: L29189; AAA42562.1; -
 DR EMBL: K00928; AAA42564.1; -
 DR PIR: A03997; VCFV37.
 DR InterPro: IPR005166; Avian_gp85.
 DR Pfam: PF03708; Avian_gp85; 1.
 DR Coat protein; Transmembrane.
 FT NON_TER 1 1
 FT CONFLICT 21 21 K -> N (IN REF. 2).
 FT CONFLICT 28 28 V -> I (IN REF. 2).
 FT CONFLICT 30 30 V -> A (IN REF. 2).
 FT CONFLICT 43 43 T -> V (IN REF. 2).
 FT CONFLICT 79 79 R -> K (IN REF. 2).
 FT CONFLICT 145 145 Q -> H (IN REF. 2).
 FT CONFLICT 169 169 L -> P (IN REF. 2).
 FT CONFLICT 212 213 ML -> IV (IN REF. 2).
 FT CONFLICT 217 217 R -> I (IN REF. 2).
 FT CONFLICT 238 238 C -> Y (IN REF. 2).
 SQ SEQUENCE 246 AA; 26991 MW; 401B95917E2FA951 CRC64;
 Query Match 5.2%; Score 7; DB 1; Length 246;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 33 ANLTSL 39
 Db 92 ANLTSL 98
 ID 11111111
 AC ENV_AVIR3 STANDARD; PRT; 257 AA.
 RX P33498;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE ENV polypeptide (coat polypeptide) [Contains: Glycoprotein 85 (Coat
 DE protein GP85); Glycoprotein 37 (Coat protein GP37)] (Fragment).
 GN ENV.
 OS Avian retrovirus RPL30.
 OC Viruses: Retroviridae; Retroviridae; Avian type C retroviruses.
 OC NCBI_TaxID=31671;
 OX 111
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92407992; PubMed-1527848;
 RA Jia R., Mayer B.J., Hanafusa T., Hanafusa H.;
 RT "A novel oncogene, v-ryk, encoding a truncated receptor tyrosine
 RT kinase is transduced into the RPL30 virus without loss of viral
 RT sequences";
 RL J. Virol. 66:5975-5987(1992).
 CC -----
 CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A ENV-RYK
 CC POLYPEPTIDE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M92847; AAA42673.1; -
 DR PIR: A43362; A43362.
 DR InterPro: IPR005166; Avian_gp85.
 DR Pfam: PF03708; Avian_gp85; 1.
 DR Coat protein; Glycoprotein; Transmembrane; Polypeptide.
 FT NON_TER 1 1
 FT CHAIN 1 78 GLYCOPROTEIN 85.
 FT CHAIN 79 257 GLYCOPROTEIN 37.
 FT TRANSMEM 228 253 POTENTIAL.
 FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 257 AA; 27915 MW; B46FC61B8F10E187 CRC64;
 Query Match 5.2%; Score 7; DB 1; Length 257;
 Best Local Similarity 100.0%; Pred. No. 6.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 33 ANLTSL 39
 Db 129 ANLTSL 135
 ID 11111111
 AC VDLG_HELPJ STANDARD; PRT; 271 AA.
 AC 092KWL;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable short-chain type dehydrogenase/reductase vldc (EC 1.-.-.-).
 GN VDLG OR JHP0823.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TaxID=85963;
 OX 111
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99120557; PubMed-9923682;
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voyis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen Helicobacter pylori";
 RT Nature 397:176-180(1999).
 RL Nature 397:176-180(1999).
 CC -----
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE001511; AAD06395.1; -
 DR HSSP: P14061; 1FDS.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 DR PRINTS: PR00080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; 1.
 KW Oxidoreductase; Complete proteome.
 NP BIND 1 25
 FT ACT_SITE 142 142
 FT ACT_SITE 142 142
 BY SIMILARITY.

SQ SEQUENCE 271 AA; 30086 MW; 23172021E1525288 CRC64;
 Query Match 5.2%; Score 7; DB 1; Length 271;
 Best Local Similarity 100.0%; Pred. No. 7.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 106 EVAKIV 112
 DB 217 EVAKIV 223
 RESULT 6
 Y149_MYCE STANDARD; PRT; 281 AA.
 AC P47395;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical lipoprotein MG149 precursor.
 GN MG149.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2097;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=96026346; PubMed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.F., Dougherty B.A., Bort K.F., Hu P.-C., Lincer T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.,
 RT "The minimal gene complement of Mycoplasma genitalium.";
 RL Science 270:397-403(1995).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 (potential).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U39696; AAC71367.1; -.
 DR TIGR; MG149; -.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Hypothetical protein; Lipoprotein; Membrane; Signal;
 KM Complete proteome.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 1 281 HYPOTHETICAL LIPOPROTEIN MG149.
 FT LIPID 24 24 N-ACYL DIGLYCERIDE (POTENTIAL).
 SO SEQUENCE 281 AA; 32428 MW; A86828BEDD005FF CRC64;
 Query Match 5.2%; Score 7; DB 1; Length 281;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 17 LASCAST 23
 DB 21 LASCAST 27
 RESULT 7
 CATO_RAT STANDARD; PRT; 343 AA.
 AC Q90ZES;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cathepsin Q precursor (EC 3.4.22.-).
 OX NCBI_TaxID=727;
 RN

GN CTSC OR CATO.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mistart; TISSUE=placenta;
 RX MEDLINE=20139729; PubMed=10673370;
 RA Sol-Church K., French J., Mason R.W.;
 RT "Cathepsin Q, a novel lysosomal cysteine protease highly expressed in
 RT placenta.";
 RL Biochem. Biophys. Res. Commun. 267:791-795(2000).
 CC -1- SUBCELLULAR LOCATION: Lysosomal (potential).
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PLACENTA.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF187323; AF01247.1; -.
 DR HSSP; P07711; ICTL.
 DR MEROPS; C01.039; -.
 DR InterPro; IPR000668; Peptidase_C1.
 DR InterPro; IPR000169; SHprot_acstle.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR PRINTS; PR00705; PAPAIN.
 DR ProDom; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; FALSE_NEG.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
 DR PROSITE; PS00640; THIOL_PROTEASE_ASN; 1.
 KW Hydrolyase; Thiol protease; Glycoprotein; Lysosome; Zymogen; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 124 ACTIVATION PEPTIDE (POTENTIAL).
 FT CHAIN 125 343 CATHEPSIN Q.
 FT ACT_SITE 149 149 BY SIMILARITY.
 FT ACT_SITE 286 286 BY SIMILARITY.
 FT ACT_SITE 310 310 BY SIMILARITY.
 FT DISULFID 146 189 BY SIMILARITY.
 FT DISULFID 180 222 BY SIMILARITY.
 FT DISULFID 280 332 BY SIMILARITY.
 FT CARBOHYD 228 228 N-LINKED (GLCNAc...) (POTENTIAL).
 FT CARBOHYD 298 298 N-LINKED (GLCNAc...) (POTENTIAL).
 SO SEQUENCE 343 AA; 39051 MW; 617A7156F3645D4 CRC64;
 Query Match 5.2%; Score 7; DB 1; Length 343;
 Best Local Similarity 100.0%; Pred. No. 8.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 27 NPKNSA 33
 DB 225 NPKNSA 231
 RESULT 8
 ABC_HAEIN STANDARD; PRT; 345 AA.
 AC P44785;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ATP-binding protein abc.
 GN ABC OR H10621.
 OS Haemophilus influenzae.
 CC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN-Rd / KM20 / ATCC 51907:
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Goeyne J.D.,
RA Scott J.D., Shirley R., Liu L.-T., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uitterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RL Rd."
RL Science 269:496-512(1995).
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See license@sib-sib.ch).
CC -----
DR EMBL, U32744; AAC22280.1; -.
DR HSSP; Q58663; 1G6H.
DR TIGR; HI0621; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transportr.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transportr; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW Transport; ATP-binding; Inner membrane; Complete proteome.
KW NP BIND
FT SEQUENCE 345 AA; 37877 MW; 7457118EB8B79C2 CRC64;
SQ
Query Match 5.2%; Score 7; DB 1; Length 345;
Best Local Similarity 100.0%; Pred. NB 8.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 67 SESAKI 73
Db 111 SESAKI 117
RESULT 9
ID SELD_HAEIN STANDARD; PRT; 348 AA.
AC P43911;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Selenide,water dikinase (EC 2.7.9.3) (Selenophosphate synthetase)
DE (Selenium donor protein).
DE SELD OR HI0200.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-Rd / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Goeyne J.D.,
RA Scott J.D., Shirley R., Liu L.-T., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uitterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RL Rd."
RL Science 269:496-512(1995).
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See license@sib-sib.ch).
CC -----
DR EMBL, U32744; AAC22280.1; -.
DR HSSP; Q58663; 1G6H.
DR TIGR; HI0621; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transportr.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transportr; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW Transport; ATP-binding; Inner membrane; Complete proteome.
KW NP BIND
FT SEQUENCE 345 AA; 37877 MW; 7457118EB8B79C2 CRC64;
SQ

```

```

RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RT Science 269:496-512(1995).
RN [2]
RN CONCEPTUAL TRANSLATION.
RA Batroch A.;
RL Unpublished observations (AUG-1995).
CC -1- FUNCTION: SYNTHESIZES SELENOPHOSPHATE FROM SELENIDE AND ATP
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + selenide + H(2)O = AMP + selenophosphate
CC + phosphate.
CC -1- COFACTOR: REQUIRES MAGNESIUM (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SELENOPHOSPHATE SYNTHETASE 1 FAMILY.
CC CLASS I SUBFAMILY.
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A UGA STOP CODON HAD TO
CC BE SKIPPED IN POSITION 16 TO PRODUCE THIS ORF. WE (REF 2) BELIEVE
CC THAT THIS RESIDUES IS MOST PROBABLY AN ACTIVE-SITE SELENOCYSTEINE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32705; AAC21869.1; ALT_FRAME.
DR TIGR: H10200;
DR InterPro: IPR000728; AIRS-related.
DR InterPro: IPR004536; SeId.
DR Pfam: PF00586; AIRS_1.
DR Pfam: PF02769; AIRS_C; 1.
DR TIGRFAMs: TIGR00476; seId; 1.
DR Transferrase; Selenium; Magnesium; ATP-binding; Selenocysteine;
KW Complete proteome.
FT ACT_SITE 16 POTENTIAL.
FT SE_CYS 16 POTENTIAL.
FT SITE 19 IMPORTANT FOR CATALYTIC ACTIVITY (BY
FT SIMILARITY).
FT NP_BIND 231 ATP (POTENTIAL).
SQ SEQUENCE 348 AA; 36523 MW; A7B1F534C57C264A CRC64;

Query Match 5.2%; Score 7; DB 1; Length 348;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 EVAKIV 112
DQ 117 EVAKIV 123

RESULT 10
CITL_MOUSE
ID CITL_MOUSE STANDARD; PRT; 352 AA.
AC Q9JJA4; Q9JKA7; Q9J7J1.
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE CysteinyI leukotriene receptor 1 (CysLTR1) (CysteinyI leukotriene D4
DE receptor) (LTD4 receptor).
OS CYSLTR1 OR CYSLTR OR CYSLTR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RX [1]
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RN STRAIN=129/SV.
RX PubMed=11226226;
RA Maekawa A., Kanooka Y., Lam B.K., Austen K.F.;
RT "Identification in mice of two isoforms of the cysteinyI leukotriene 1

```



```

DR PIR: A28437; DCBYO.
DR HSSP: P07805; 10U4.
DR SGD: S0001667; SPEI.
DR InterPro: IPR000183; Decarboxylase.
DR Pfam: PF002784; Orn_DAP_Arx_dec; 1.
DR Pfam: PF02784; Orn_Arx_dec; 1.
DR PRINTS: PR01179; ODADCRBLASE.
DR PROSITE: PS00878; ODR_DC_2-1; 1.
DR PROSITE: PS00879; ODR_DC_2-2; 1.
DR Lyase: Decarboxylase; Pyridoxal phosphate; Polyamine biosynthesis.
FT BINDING 116 116 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT ACT_SITE 411 411 BY SIMILARITY.
SQ SEQUENCE 466 AA; 52285 MW; 2EBD/EEBCFAA67CD CRC64;

Query Match 5.2%; Score 7; DB 1; Length 466;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 TAFTLAS 19
| | | | |
Db 326 TAFTLAS 332

RESULT 12
DNA_BORBU STANDARD; PRT; 486 AA.
ID DNA_BORBU
AC P33768;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chromosomal replication initiator protein dnaA.
GN DNA OR BB0437.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID-139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-212;
RX MEDLINE-9336157; PubMed-8359672;
RA Old I.G., Margarita D., Saint-Girons I.;
RT "Unique genetic arrangement in the dnaA region of the Borrelia
  burgdorferi linear chromosome: nucleotide sequence of the dnaA
  gene."
RL FEMS Microbiol. Lett. 111:109-114(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / B31;
RX MEDLINE-98065943; PubMed-9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
  Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
  Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
  Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
  van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weiman J.,
  Uterback T., Watney L., McDonald L., Atlach P., Bowman C.,
  Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
  Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
  burgdorferi."
RL Nature 390:580-586(1997).
RN [3]
RP FUNCTION: PLAYS AN IMPORTANT ROLE IN THE INITIATION AND REGULATION
  OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN OF REPLICATION;
  IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA AT A 9 BP CONSENSUS
  (DNAA BOX): 5'-TTATC(C/A)A(C/G)A-3'. DNAA BINDS TO ATP AND TO
  ACIDIC PHOSPHOLIPIDS (BY SIMILARITY).
RN [4]
RP SIMILARITY: BELONGS TO THE DNAA FAMILY.
RN [5]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).

```

```

CC EMBL: U04527; AA58941.1; -.
DR EMBL: AE001149; AAB91515.1; -.
DR TIGR: BB0437; -.
DR InterPro: IPR001957; Bac_DnaA.
DR Pfam: PF00308; bac_dnaA; 1.
DR PRINTS: PR00051; DnaA.
DR TIGRfam: TIGR00362; DnaA; 1.
DR PROSITE: PS01008; DnaA; 1.
DR DNA replication; DNA-binding; ATP-binding; Complete proteome.
FT NP_BIND 183 190 ATP (POTENTIAL).
SQ SEQUENCE 486 AA; 56783 MW; 27CA6BE6A9BA9BC CRC64;

Query Match 5.2%; Score 7; DB 1; Length 486;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 KILYVTA 9
| | | | |
Db 208 KILYVTA 214

RESULT 13
ENV_RSVP STANDARD; PRT; 603 AA.
ID ENV_RSVP
AC P03396;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ENV polypeptide (Contains: Surface protein GP85, Membrane protein
  GP37).
GN ENV.
OS Rous sarcoma virus (strain Prague C).
OC Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.
OX NCBI_TaxID-11888;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-83155662; PubMed-6299578;
RA Schwartz D., Tizard R., Gilbert W.;
RT "Nucleotide sequence of Rous sarcoma virus."
RL Cell 32:853-869(1983).
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: AS A RESULT OF BASE VARIATIONS, A DIFFERENT VERSION
  OF THIS SEQUENCE MAY EXIST HAVING 17-GLU, 134-SER, 158-SER,
  334-THR, 383-THR, 392-VAL, 522-GLU, 541-LEU, AND 567-VAL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL: V01197; -. NOT ANNOTATED CDS.
DR PIR: A03996; VCFEVR.
DR InterPro: IPR005166; Avian_gp85.
DR Pfam: PF03708; Avian_gp85; 1.
RN [4]
RP Coat protein; Polyprotein; Transmembrane; Signal.
FT SIGNAL 1 64
FT CHAIN 65 603 SURFACE PROTEIN GP85.
FT CHAIN 406 603 MEMBRANE PROTEIN GP37.
FT TRANSMEM 557 577 POTENTIAL.
SQ SEQUENCE 603 AA; 65660 MW; D44EC2AA62282C94 CRC64;

Query Match 5.2%; Score 7; DB 1; Length 603;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39
| | | | |
Db 456 ANLTSL 462

```



```

RESULT 14
MNT3_YEAST
ID MNT3_YEAST STANDARD: PRT: 630 AA.
AC P40549:
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha-1,3-mannosyltransferase MNT3 (EC 2.4.1.-).
GN MNT3 OR Y1014W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Barrett B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA Genes S., Hamlyn N., Horsnell T.S., Hunt S., Javelin K., Jones M.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Mannosyltransferase involved in adding the 4th and 5th
CC mannose residues of O-linked glycans.
CC -1- PATHWAY: Glycosylation.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential).
CC -1- SIMILARITY: BELONGS TO THE MNT1/MNT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL: 246881; CAAB6979.1; -
DR SCD: S0001276; MNT3.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 14 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 15 31 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT (POTENTIAL).
FT DOMAIN 32 630 LUMENAL (POTENTIAL).
FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 630 AA: 72409 MW: 68087/DD003D8CD91 CRC64;

Query Match 5.2%; Score 7; DB 1; Length 630;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 AKISETA 77
DB 280 AKISETA 286

```

```

RESULT 15
DPL1_HUMAN
ID DPL1_HUMAN STANDARD: PRT: 977 AA.
AC O14490; O14489; P78335;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Disks large-associated protein 1 (DAP-1) (Guanylate kinase-associated
DE protein) (HGAP) (SAP90/PSD-95-associated protein 1) (SAPAP1) (PSD-
DE 95/SAP90 binding protein 1).
GN DLGAP1 OR GKAP OR DAP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Brain;
RX MEDLINE=97431353; Pubmed=9286858;
RA Satoh K., Yanai H., Senda T., Kohu K., Nakamura T., Okumura N.,
RA Matsumine A., Kobayashi S., Toyoshima K., Akiyama T.;
RT "DAP-1, a novel protein that interacts with the guanylate kinase-like
RT domains of hDLG and PSD-95."
RL Genes Cells 2:415-424(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RX MEDLINE=97177144; Pubmed=9024696;
RA Kim E., Naisbitt S., Hsueh Y.-P., Rao A., Rothschild A., Craig A.M.,
RA Sheng M.;
RT "GKAP, a novel synaptic protein that interacts with the guanylate
RT kinase-like domain of the PSD-95/SAP90 family of channel clustering
RT molecules."
RL J. Cell Biol. 136:669-678(1997).
CC -1- FUNCTION: May play a role in the molecular organization of
CC synapses and neuronal cell signaling. Could be an adapter protein
CC linking ion channel to the subsynaptic cytoskeleton. May induce
CC enrichment of PSD-95/SAP90 at the plasma membrane.
CC -1- SUBUNIT: Interacts with DLG1, DLG4/PSD-95 and APC (By similarity).
CC Isoform 1 and isoform 2 C-terminus bind the PDZ domain of SHANK1,
CC SHANK2 and SHANK3 (By similarity). Is part of a complex with
CC DLG4/PSD-95 and SHANK1, SHANK2 or SHANK3 (By similarity).
CC -1- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; 1/DAP1-alpha (shown here),
CC 2/DAP1-beta and 3; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: Expressed in brain.
CC -1- SIMILARITY: BELONGS TO THE SAPAP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL: AB000277; BAA23258.1; -
DR EMBL: AB000276; BAA23257.1; -
DR EMBL: U67988; AAC51119.1; -
DR Genew; HGNC:2905; DLGAP1.
DR MIM: 605445; -
DR InterPro: IPR005026; GKAP.
DR Pfam: PF03359; GKAP. 1.
KW Membrane; Alternative splicing.
FT DOMAIN 1 622 630 POLY-THR.
FT VARSPLIC 1 302 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT VARSPLIC 303 319 MYKESCOOESCOYLO -> MNLIFHMDLFGPANK
FT FT (IN ISOFORM 2 AND ISOFORM 3).
FT VARSPLIC 909 929 ERRAPPPYKAPKAPAPLIR -> VEGCRCPVHLYATCTN
FT FT (IN ISOFORM 3).
FT VARSPLIC 930 977 TGQSK (IN ISOFORM 3).
FT CONFLICT 734 734 A -> P (IN REF. 2).
FT CONFLICT 740 740 S -> T (IN REF. 2).
FT CONFLICT 752 753 AA -> SP (IN REF. 2).
SQ SEQUENCE 977 AA: 108872 MW: 670F72B17D9BE67 CRC64;

Query Match 5.2%; Score 7; DB 1; Length 977;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 STLETTT 96
DB 383 STLETTT 389

```

```

RESULT 16
DPL1_RAT
ID DPL1_RAT STANDARD: PRT: 992 AA.

```


AC P97836: P97841: 054773;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Disks large-associated protein 1 (DAP-1) (Guanylate kinase-associated protein) (GKAP) (SAP90/PSD-95-associated protein 1) (SAPAP1) (PSD-95/SAP90 binding protein 1).
 GN DLGAP1 OR GKAP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain;
 RA MEDLINE=97177144; PubMed=9024696;
 RA Kim E., Natsubitt S., Hsueh Y.-P., Rao A., Rothschild A., Craig A.M., Sheng M.;
 RA "GKAP, a novel synaptic protein that interacts with the guanylate kinase-like domain of the PSD-95/SAP90 family of channel clustering molecules.";
 RT J. Cell Biol. 136:669-678(1997).
 RL [2]
 RN RP SEQUENCE FROM N.A. (ISOFORM 3).
 RC MEDLINE=98089008; PubMed=9428732;
 RA Kawashima N., Takamiya K., Sun J., Kitabatake A., Sobue K.;
 RA "Differential expression of isoforms of PSD-95 binding protein (GKAP/SAPAP1) during rat brain development.";
 RT FEBS Lett. 418:301-304(1997).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain;
 RA MEDLINE=97277335; PubMed=9115257;
 RA Takeuchi M., Hata Y., Hirao K., Toyoda A., Irie M., Takai Y.;
 RA "SAPAPs, a family of PSD-95/SAP90-associated proteins localized at postsynaptic density.";
 RT J. Biol. Chem. 272:11943-11951(1997).
 RL [4]
 RN RP INTERACTION WITH DLG4 AND SHANK PROTEINS.
 RA PubMed=10527873;
 RA Boeckers T.M., Winter C., Smalla K.-H., Kreutz M.R., Bockmann J., Seidenbecher C., Garner C.C., Gundelfinger E.D.;
 RA "Proline-rich synapse-associated proteins Prosap1 and Prosap2 interact with synaptic proteins of the SAPAP/GKAP family.";
 RT Biochem. Biophys. Res. Commun. 264:247-252(1999).
 RL [5]
 RN RP INTERACTION WITH DLG4 AND SHANK1.
 RC TISSUE=Brain;
 RA MEDLINE=99419021; PubMed=10488079;
 RA Yao I., Hata Y., Hirao K., Deguchi M., Ide N., Takeuchi M., Takai Y.;
 RA "Synamon, a novel neuronal protein interacting with synapse-associated protein 90/Postsynaptic density-95-associated protein.";
 RT J. Biol. Chem. 274:27463-27466(1999).
 RL [6]
 RN RP INTERACTION WITH SHANK PROTEINS, AND MUTAGENESIS ALA-990 AND ALA-992.
 RA MEDLINE=99360650; PubMed=10433268;
 RA Natsubitt S., Kim E., Tu J.C., Xiao B., Sala C., Valtchanoff J., Weinberg R.J., Worley P.F., Sheng M.;
 RA "Shank, a novel family of postsynaptic density proteins that binds to the NMDA receptor/PSD-95/GKAP complex and cortactin.";
 RT Neuron 23:569-582(1999).
 RL [7]
 RN RP FUNCTION: May play a role in the molecular organization of synapses and neuronal cell signaling. Could be an adapter protein linking ion channel to the subsynaptic cytoskeleton. May induce enrichment of PSD-95/SAP90 at the plasma membrane.
 CC [1-] SUBUNIT: Interacts with DLG1, DLG4/PSD-95 and APC. The C-terminus binds the PDZ domain of the SHANK1, SHANK2 and SHANK3.
 CC [1-] SUBCELLULAR LOCATION: Membrane-associated.
 CC [1-] ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3; are produced by alternative splicing.
 CC [1-] TISSUE SPECIFICITY: Expressed in brain and testis.
 CC [1-] SIMILARITY: BELONGS TO THE SAPAP FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U67987; AAC53054.1; -;
 DR EMBL: AB003594; BAA24265.1; -;
 DR EMBL: U67137; AAB48587.1; -;
 DR InterPro: IPR005026; GKAP.
 DR Pfam: PF03359; GKAP; 1.
 KW Membrane; Alternative splicing.
 FT VARSPPLIC 1 298 MISSING (IN ISOFORM 2 AND ISOFORM 3).
 FT VARSPPLIC 299 324 OKASVNDQAVVSEACQDQSCQYL -> MIDLFRAEWS
 FT VARSPPLIC 325 325 O -> QVW (IN ISOFORM 2).
 FT VARSPPLIC 547 574 MISSING (IN ISOFORM 2).
 FT MUTAGEN 990 990 T->A: ABOLISHES INTERACTION WITH SHANK1.
 FT MUTAGEN 992 992 L->A: ABOLISHES INTERACTION WITH SHANK1.
 FT CONFLICT 636 636 S -> T (IN REF. 1 AND 2).
 SQ SEQUENCE 992 AA; 110177 MW; 44BAF9BC0C14C099 CRC64;
 Query Match 5.28; Score 7; DB 1; Length 992;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 90 SLELETT 96
 Db 389 SLELETT 395
 RESULT 17
 RL32_PASMU STANDARD; PRT; 55 AA.
 ID 09CIS9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 50S ribosomal protein L32.
 GN RPLP6 OR RPL32 OR PM1912.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Pm70;
 RA MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitlam T.S., Kapur V.;
 RA "Complete genomic sequence of Pasteurella multocida Pm70.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 RL [1-] SIMILARITY: BELONGS TO THE L32P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AE006228; AAK03996.1; -;
 DR InterPro: IPR002677; Ribosomal_L32p.
 DR Pfam: PF01783; Ribosomal_L32p; 1.
 DR TIGRFAMs: TIGR01031; rplm_Pdact; 1.
 KW Ribosomal protein; Complete proteome.
 FT INIT_MET 0 0 BY SIMILARITY.
 SQ SEQUENCE 55 AA; 6302 MW; 2B45A37A32A5A239 CRC64;

```

Query Match
Best Local Similarity 4.58; Score 6; DB 1; Length 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 94 LTTAAI 99
    |||||
Db 21 LTTAAI 26

RESULT 18
Y147_UREPA STANDARD; PRT; 57 AA.
AC 09P027;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein UUI47.
GN UUI47.
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
OX NCBI_TaxID=134821;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=Sevovar 3;
RX MEDLINE=20500219; PubMed=11048724;
RA Glass J.I., Letkowitz E.J., Glass J.S., Helner C.R., Chen E.Y.,
RA Cassell G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
RT urealyticum."
CC -----
CC Nature 407:757-762(2000).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE002115; AAF30553.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 57 AA; 6987 MW; 9A72EC5D476B3A8 CRC64;

Query Match
Best Local Similarity 4.58; Score 6; DB 1; Length 57;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 89 VSLTEL 94
    |||||
Db 11 VSLTEL 16

RESULT 19
SECE_SULAC STANDARD; PRT; 65 AA.
AC P27340;
DT 01-NOV-1992 (Rel. 23, Created)
DT 01-NOV-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Preprotein translocase sece subunit (Protein transport protein SECE1
DE gamma subunit homolog).
GN SECE.
OS Sulfolobus acidocaldarius.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2285;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=92048486; PubMed=1658539;
RA Ramirez C., Matheson A.T.;
RT "A gene in the archaeobacterium Sulfolobus solfataricus that codes for
RT a protein equivalent to the alpha subunits of the signal recognition
RT particle receptor in eukaryotes."

```

```

RL MOL. Microbiol. 5:1687-1693(1991).
RN 121
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33909 / NCIB 11770 / DSM 639;
RX MEDLINE=95226466; PubMed=7711082;
RA Moll R., Schmidtke S., Schaefer G.;
RT "Nucleotide sequence of a gene cluster encoding ribosomal proteins in
RT the thermoacidophilic crenarchaeon Sulfolobus acidocaldarius."
RL Biochim. Biophys. Acta 1261:315-318(1995).
CC -1- FUNCTION: Involved in protein export (By similarity).
CC -1- SUBUNIT: Component of the protein translocase complex (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Tail-anchored membrane protein (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE SECE/SEC61-GAMMA FAMILY.
CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO ORIGINATE FROM
CC S.SOLFATARICUS STRAIN P1, BUT THE CULTURE WAS CONTAMINATED WITH
CC S.ACIDOCALDARIUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X58538; CAA41430.1; -
DR EMBL; X77509; CAA34644.1; -
DR PIR; S16483; S16483.
DR PIR; S41969; S41969.
DR InterPro: IPR001901; SECE.
DR InterPro: IPR004795; SECE_euk_arch.
DR TRIGRAMS: TIGR00327; SECE_euk_arch. 1.
DR PROSITE: PS01067; SECE_SEC61G; 1.
KW Protein transport; Translocation; Transmembrane.
FT TRANSMEM 44
FT POTENTIAL 64
SQ SEQUENCE 65 AA; 7553 MW; 62B2218173852809 CRC64;

Query Match
Best Local Similarity 4.58; Score 6; DB 1; Length 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 35 LTTSLI 40
    |||||
Db 58 LTTSLI 63

RESULT 20
RS15_SYNY3 STANDARD; PRT; 89 AA.
AC P72866;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 30S ribosomal protein S15.
GN RPSO OR RPS15 OR S15L784.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA MiyaJima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
RA Okumura S., Shimo S., Takeuchi C., Wada T., Watanabe A.,
RA Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions."
RL DNA Res. 3:109-136(1996).
CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE 16S RIBOSOMAL RNA BINDING

```

CC PROTEINS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE S1SP FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D90901; BAA16882.1; -
 CC HSP: P80378; IAB3.
 CC InterPro: IPR005290; RS15_bact.
 CC InterPro: IPR000589; Ribosomal_S15.
 CC Pfam: PF00312; Ribosomal_S15; 1.
 CC TIGRfam: TIGR00952; S15_bact; 1.
 CC PROSITE: PS00362; RIBOSOMAL_S15; 1.
 CC RIBOSOMAL PROTEIN; rRNA-binding; Complete proteome.
 CC KW RIBOSOMAL PROTEIN; rRNA-binding; A22EC821584C556 CRC64;
 CC SO SEQUENCE 89 AA; 10373 MW; 422EC821584C556 CRC64;
 CC
 CC Query Match 4.5%; Score 6; DB 1; Length 89;
 CC Best Local Similarity 100.0%; Pred. No. 31;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 50 TQLTGH 55
 CC | | | | |
 CC Db 37 TQLTGH 42
 CC
 CC RESULT 21
 CC MPIL_LYMST STANDARD; PRT; 123 AA.
 CC AC P07223; O9TWM5; O9TWM6;
 CC DT 01-APR-1988 (Rel. 07, Created)
 CC DT 01-MAR-1989 (Rel. 10, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Molluscan insulin-related peptide 1 precursor (MIP 1).
 CC OS Lymnaea stagnalis (Great pond snail).
 CC CC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
 CC CC Lymnaeidae; Lymnaea.
 CC CC NCBI_TaxID=6523;
 CC RN [1]
 CC RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 CC RX MEDLINE=88122643; PubMed=3340203;
 CC RA Smit A.B., Vreugdenhil E., Eberink R.H.M., Geraerts W.P.M.,
 CC RA Klootwijk J., Joosse J.,
 CC RT "Growth-controlling molluscan neurons produce the precursor of an
 CC RT insulin-related peptide."
 CC RL Nature 331:535-538(1988).
 CC RN [2]
 CC RP SEQUENCE OF 32-67; 72-96 AND 99-123.
 CC RC TISSUE-Light-green cells;
 CC RX MEDLINE=92405808; PubMed=1526314;
 CC RA Li K.W., Geraerts W.P.M., Eberink R.H.M., Joosse J.,
 CC RT "Purification and sequencing of molluscan insulin-related peptide I
 CC RT (MIP I) from the neuroendocrine light green cells of Lymnaea
 CC RT stagnalis."
 CC RL Mol. Cell. Endocrinol. 85:141-150(1992).
 CC RL -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
 CC DISULFIDE BONDS.
 CC -1- SUBCELLULAR LOCATION: Secretory granules.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE CEREBRAL LIGHT-GREEN CELLS
 CC WHICH ARE GIANT NEUROENDOCRINE CELLS INVOLVED IN THE CONTROL OF
 CC GROWTH.
 CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X06983; CAA30043.1; -
 CC DR PIR: S00531; S00531.
 CC DR InterPro: IPR004825; Ins/IGF/relax.
 CC DR InterPro: IPR003234; Mollusc_ins.
 CC DR Pfam: PF00049; Insulin; 1.
 CC DR PRIMS: PR00276; INSULIN.
 CC DR ProDom: PD015667; Mollusc_ins; 1.
 CC DR SMART: SM00078; IIGF; 1.
 CC DR PROSITE: PS00262; INSULIN; 1.
 CC KW Signal; Cleavage on pair of basic residues; Insulin family;
 CC KW Multigene family.
 CC FT CHAIN 1 31
 CC FT SIGNAL 32 67
 CC FT PROPEP 72 96
 CC FT CHAIN 99 123
 CC FT MOD_RES 32 32
 CC FT MOD_RES 99 99
 CC FT DISULFID 48 109
 CC FT DISULFID 60 122
 CC FT DISULFID 108 113
 CC SO SEQUENCE 123 AA; 13458 MW; 87C89F042E2D700B CRC64;
 CC
 CC Query Match 4.5%; Score 6; DB 1; Length 123;
 CC Best Local Similarity 100.0%; Pred. No. 42;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 64 KLSSES 69
 CC | | | | |
 CC Db 85 KLSSES 90
 CC
 CC RESULT 22
 CC YCVA_YEAST STANDARD; PRT; 136 AA.
 CC AC P25640;
 CC DT 01-MAY-1992 (Rel. 22, Created)
 CC DT 01-MAY-1992 (Rel. 22, Last sequence update)
 CC DT 15-DEC-1998 (Rel. 37, Last annotation update)
 CC DE Hypothetical 15.8 kDa protein in SUP61-RAD18 intergenic region.
 CC GN YCR064C OR YCR64C.
 CC OS Saccharomyces cerevisiae (Baker's yeast).
 CC CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 CC CC NCBI_TaxID=4932;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RA Antoine G., Benit P., Chanet R., Fabre R., Faye G., Fukuhara H.,
 CC RA Mathieu A., Sor F.,
 CC RL Submitted (MAR-1992) to the EMBL/Genbank/DBJ databases.
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X59720; CAA42279.1; -
 CC DR PIR: S19479; S19479.
 CC DR SCD: S0000660; YCR064C.
 CC KW Hypothetical protein.
 CC SO SEQUENCE 136 AA; 15776 MW; 0C5A14D207D78FD6 CRC64;
 CC
 CC Query Match 4.5%; Score 6; DB 1; Length 136;
 CC Best Local Similarity 100.0%; Pred. No. 46;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 129 VPTTTR 134

Db 118 VPTTTR 123

|||||

RESULT 23

MP12_LYMST STANDARD; PRT; 137 AA.

AC P25289;

DT 01-MAY-1992 (Rel. 22, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

DE Molluscan insulin-related peptide 2 precursor (MIP II).

OS Lymnaea stagnalis (Great pond snail).

OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;

OC Lymnaeidae; Lymnaea.

OX NCBI_TaxID=6523;

RN (1)

RP SEQUENCE FROM N.A.

RC TISSUE=CNS;

RX MEDLINE=91330924; PubMed=1868853;

RA Smit A.B., Geraerts W.P.M., Meester I., van Heerikhuisen H., Joosse J.;

RT "Characterization of a cDNA clone encoding molluscan insulin-related peptide II of Lymnaea stagnalis.";

RL Eur. J. Biochem. 199;699-703(1991).

RN (2)

RP SEQUENCE OF 32-66 AND 113-137.

RC TISSUE=Light-green cells;

RX MEDLINE=92283200; PubMed=1350761;

RA Li K.-W., Geraerts W.P.M., Joosse J.;

RT "Purification and sequencing of molluscan insulin-related peptide II from the neuroendocrine light green cells in Lymnaea stagnalis.";

RL Endocrinology 130:3427-3432(1992).

CC -1 SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO DISULFIDE BONDS.

CC -1 SUBCELLULAR LOCATION: Secretory granules.

CC -1 TISSUE SPECIFICITY: EXPRESSED IN THE CEREBRAL LIGHT-GREEN CELLS WHICH ARE GIANT NEUROENDOCRINES CELLS INVOLVED IN THE CONTROL OF GROWTH.

CC -1 SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

CC EMBL; X59302; CAA41989.1; -.

DR PIR; S17195; S17195.

DR InterPro; IPR004825; Ins/IGF/relax.

DR InterPro; IPR003234; Mollusc_ins.

DR Pfam; PF00049; Insulin; 1.

DR ProDom; PD015667; Mollusc_ins; 1.

DR SMART; SM00078; IIGF; 1.

DR PROSITE; PS00262; INSULIN; 1.

DR Signal; Cleavage on pair of basic residues; Insulin family; Multigene family.

KW Signal; Cleavage on pair of basic residues; Insulin family; Multigene family.

FT SIGNAL 1 31

FT CHAIN 32 68

FT PROPEP 71 83

FT PROPEP 86 110

FT CHAIN 113 137

FT MOD_RES 32 32

FT MOD_RES 113 113

FT DISULFID 47 123

FT DISULFID 59 136

FT DISULFID 122 127

FT DISULFID 137 AA; 15257 MW; B9E56409D1ECB41 CRC64;

Query Match 4.5%; Score 6; DB 1; Length 137;

Best local Similarity 100.0%; Pred. No. 46;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 KLSSES 69

Db 99 KLSSES 104

|||||

RESULT 24

MP15_LYMST STANDARD; PRT; 138 AA.

AC P31241;

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

DE Molluscan insulin-related peptide 5 precursor (MIP V).

OS Lymnaea stagnalis (Great pond snail).

OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;

OC Lymnaeidae; Lymnaea.

OX NCBI_TaxID=6523;

RN (1)

RP SEQUENCE FROM N.A.

RC TISSUE=CNS;

RX MEDLINE=92356799; PubMed=1323019;

RA Smit A.B., Thijssen S.F.T., Geraerts W.P.M., Meester I., Heerikhuisen H., Joosse J.;

RT "Characterization of a cDNA clone encoding molluscan insulin-related peptide V of Lymnaea stagnalis.";

RL Brain Res. Mol. Brain Res. 14:7-12(1992).

CC -1 SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO DISULFIDE BONDS.

CC -1 SUBCELLULAR LOCATION: Secretory granules.

CC -1 TISSUE SPECIFICITY: EXPRESSED IN THE CEREBRAL LIGHT-GREEN CELLS WHICH ARE GIANT NEUROENDOCRINES CELLS INVOLVED IN THE CONTROL OF GROWTH.

CC -1 SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

CC EMBL; X59302; -; NOT_ANNOTATED_CDS.

DR PIR; A43957; A43957.

DR InterPro; IPR004825; Ins/IGF/relax.

DR InterPro; IPR003234; Mollusc_ins.

DR Pfam; PF00049; Insulin; 1.

DR ProDom; PD015667; Mollusc_ins; 1.

DR SMART; SM00078; IIGF; 1.

DR PROSITE; PS00262; INSULIN; 1.

DR Signal; Cleavage on pair of basic residues; Insulin family; Multigene family.

KW Signal; Cleavage on pair of basic residues; Insulin family; Multigene family.

FT SIGNAL 1 31

FT CHAIN 32 69

FT PROPEP 72 84

FT PROPEP 87 111

FT CHAIN 114 138

FT MOD_RES 32 32

FT MOD_RES 114 114

FT DISULFID 48 124

FT DISULFID 60 137

FT DISULFID 123 128

FT DISULFID 138 AA; 15540 MW; 9EF1921B84A9D6CA CRC64;

Query Match 4.5%; Score 6; DB 1; Length 138;

Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 KLSSES 69
|||||
DB 100 KLSSES 105

RESULT 25

PSB2_ECOLI STANDARD; PRT; 144 AA.

ID PSB2_ECOLI
AC P10032:
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE PSIB protein.
GN PSIB.
OS Escherichia coli.
OC Plasmid R6-5.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]

SEQUENCE FROM N.A.

RA MEDLINE=89083490; PubMed=3205720;
RA Dutreix M., Baeckman A., Celerier J., Bagdasarian M.M., Sommer S.,
RA Ballone A., Devoret R., Bagdasarian M.;
RT "Identification of psib genes of plasmids R-6 and R-6-5. Molecular basis
for psib enhanced expression in plasmid R-6-5."
RL Nucleic Acids Res. 16:10669-10679(1988).
CC -1- FUNCTION: INHIBITION OF THE SOS PATHWAY.
CC -1- SIMILARITY: STRONG, TO PSIB PROTEIN OF PLASMID F.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)
CC -----

DR EMBL; X12463; CA831004.1; -.
DR PIR; S01897; S01897.
KW Plasmid.

SEQUENCE 144 AA; 15777 MW; 369A28FB96B0887 CRC64;

Query Match 4.5%; Score 6; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 48;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 85 APEAVS 90
|||||
DB 109 APEAVS 114

RESULT 26

RS13_YEAST STANDARD; PRT; 150 AA.

ID RS13_YEAST
AC P05756:
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 40S ribosomal protein S13 (S27A) (YS15).
GN RPS13 OR RPS13C OR YDR064W OR YD9609.18 OR D4252.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]

RA SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RA Hunt S., Bowman S., Harris D., Barrell B.G., Rajandream M.A.;
Submitted (May-1995) to the EMBL/GenBank/DBJ databases.

RN [2]
RP SEQUENCE FROM N.A.

RX MEDLINE=96381250; PubMed=8789263;
RA Brandt P., Ramlov S., Otto B., Bloeker H.;
RT "Nucleotide sequence analysis of a 32,500 bp region of the right arm
of Saccharomyces cerevisiae chromosome IV."
RL Yeast 12:85-90(1996).
RN [3]

RP SEQUENCE OF 1-40.
RX MEDLINE=83048950; PubMed=6814480;
RA Olaka E., Higo K.-I., Osawa S.;

RT "Isolation of seventeen proteins and amino-terminal amino acid
sequences of eight proteins from cytoplasmic ribosomes of yeast."
RL Biochemistry 21:4545-4550(1982).
CC -1- SIMILARITY: BELONGS TO THE S15P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)
CC -----

DR EMBL; 249209; CAA89093.1; -.
DR EMBL; X84162; CAA58980.1; -.
DR EMBL; 274360; CAA88882.1; -.
DR PIR; S11578; S11578.
DR SGD; S0002471; RPS13.
DR InterPro; IPR000589; Ribosomal_S15.
DR Pfam; PF00312; Ribosomal_S15; 1.
DR PROSITE; PS00362; RIBOSOMAL_S15; 1.
KW Ribosomal protein.

FT INT_MEN 0
FT CONFLICT 24 W -> G (IN REF. 3).
FT CONFLICT 31 S -> C (IN REF. 3).
SQ SEQUENCE 150 AA; 16898 MW; 378D72BE81E3AF58 CRC64;

Query Match 4.5%; Score 6; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 KLSSES 69
|||||
DB 26 KLSSES 31

RESULT 27

MPAG_ALINGL STANDARD; PRT; 159 AA.

ID MPAG_ALINGL
AC P38948:
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major pollen allergen Aln g 1 (Aln g I).
OS Alnus glutinosa (Alder).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids I; Fagales; Betulaceae; Alnus.
OX NCBI_TaxID=3517;
RN [1]

SEQUENCE FROM N.A.

RA TISSUE=Pollen;
RX MEDLINE=93094476; PubMed=1460197;
RA Breiteneder H., Ferreira F., Reikertstorfer A., Duchene M.,
RA Valenta R., Hofmann-Sommergruber K., Ebner C., Breitenbach M.,
RA Kraft D., Scheiner O.;

RT "Complementary DNA cloning and expression in Escherichia coli of Aln
g 1, the major allergen in pollen of alder (Alnus glutinosa).";
RL J. Allergy Clin. Immunol. 90:909-917(1992).
CC -1- DISEASE: A CAUSE OF TYPE I ALLERGIC REACTIONS IN EUROPE, NORTH
CC AMERICA AND USSR.

CC -1- SIMILARITY: BELONGS TO THE BETVI FAMILY OF PATHOGENESIS-RELATED


```

CC -1- SUBUNIT: PART OF A COMPLEX THAT CONTAINS SEC61, SEC62, SEC63,
CC SEC66 AND SEC72.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L29340; AAA21840.1; -
DR EMBL: L07116; -; NOT_ANNOTATED_CDS.
DR EMBL: U17243; AAB67338.1; -
DR EMBL: L08690; -; NOT_ANNOTATED_CDS.
DR PIR: A53835; A53835
DR SCD: S0004283; SEC72.
DR InterPro: IPR001440; TPR.
DR Pfam: PF00515; TPR; 2.
DR Transport: Protein transport.
FT INIT_MET 0 0 V -> L (IN REF. 1).
FT CONFLICT 32 32 L -> M (IN REF. 1).
FT CONFLICT 186 186 M; 757DC9492D6571BD CRC64;
SQ SEQUENCE 192 AA; 21476 MW; 757DC9492D6571BD CRC64;

Query Match 4.5%; Score 6; DB 1; Length 192;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 35 LTTSLI 40
Db 33 LTTSLI 38
|||||

RESULT 30
RETB_CHICK STANDARD; PRT; 196 AA.
ID RETB_CHICK
AC P41263;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasma retinol-binding protein precursor (PRBP).
GN RBP4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95267350; PubMed=7748490;
RA Vieira A.V., Kuchler K., Schneider W.J.;
RT "Retinol in avian oogenesis: molecular properties of the carrier
RT protein".
RL DNA Cell Biol. 14:403-410(1995).
CC -1- FUNCTION: Delivers retinol from the liver stores to the peripheral
CC tissues. In plasma, the RBP-retinol complex interacts with
CC transthyretin, this prevents its loss by filtration through the
CC kidney glomeruli (by similarity).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X77960; CAA54922.1; -
DR PIR: S42887; S42887.

```

```

DR HSSP: P02753; IRBP.
DR InterPro: IPR002345; Lipocalin.
DR InterPro: IPR000566; Lipocalin_cytrFABP.
DR Pfam: PF00061; Lipocalin_1.
DR PRINTS: PR00179; LIPOCALIN.
DR PROSITE: PS00213; LIPOCALIN; 1.
DR Plasma; Vitamin A; Retinol-binding; Transport; Liver; Signal;
KW Lipocalin.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 196 PLASMA RETINOL-BINDING PROTEIN.
FT DISULFID 25 181 BY SIMILARITY.
FT DISULFID 91 195 BY SIMILARITY.
FT DISULFID 141 150 BY SIMILARITY.
SQ SEQUENCE 196 AA; 22515 MW; 5E9423A14578DA75 CRC64;

Query Match 4.5%; Score 6; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 108 AOKIVR 113
Db 169 AOKIVR 174
|||||

RESULT 31
RETB_HUMAN STANDARD; PRT; 201 AA.
ID RETB_HUMAN
AC P02753; O43478; O43479; Q8MWA3; Q9P178;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasma retinol-binding protein precursor (PRBP) (RBP) (PRO2222).
GN RBP4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84069802; PubMed=6316270;
RA Colantuoni V., Romano V., Bensl G., Santoro C., Costanzo F.,
RA Raugel G., Cortese R.;
RT "Cloning and sequencing of a full length cDNA coding for human
RT retinol-binding protein".
RL Nucleic Acids Res. 11:7769-7776(1983).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-189 FROM N.A.
RX MEDLINE=86055755; PubMed=2998779;
RA D'Onofrio C., Colantuoni V., Cortese R.;
RT "Structure and cell-specific expression of a cloned human retinol
RT binding protein gene: the 5'-flanking region contains hepatoma
RL specific transcriptional signals."
EMBO J. 4:1981-1989(1985).
RN [4]
RP SEQUENCE OF 19-201, AND DISULFIDE BONDS.
RX MEDLINE=88019004; PubMed=2444024.
RA Rask L., Anundi H., Fohlman J., Peterson P.A.;
RT "The complete amino acid sequence of human serum retinol-binding
RT protein".
RL Ups. J. Med. Sci. 92:115-146(1987).
RN [5]
RP SEQUENCE OF 19-201.
RX MEDLINE=81254137; PubMed=6942701;
RA Rask L., Anundi H., Boehme J., Eriksson U., Ronne H., Sege K.,
RA Peterson P.A.;
RT "Structural and functional studies of vitamin A-binding proteins.";
RL Ann. N.Y. Acad. Sci. 359:79-90(1981).
RN [6]

```


RP SEQUENCE OF 19-183.
 RX MEDLINE=80004132; PubMed=573217;
 RA Rask L., Anundt H., Peterson P.A.;
 RT "The primary structure of the human retinol-binding protein.";
 RL FEBS Lett. 104:55-58(1979).
 RN [17]
 RP SEQUENCE OF 18-201 FROM N.A.
 RC TISSUE=Liver;
 RA Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,
 RT He F.;
 RL "Functional prediction of the coding sequences of 79 new genes deduced
 by analysis of cDNA clones from human fetal liver.";
 RN Submitted (JAN-1999) to the EMBL/GenBank/DBD databases.
 RP [8]
 RP PARTIAL SEQUENCE.
 RX MEDLINE=95395382; PubMed=7666002;
 RA Jacot S., Rose K., Hughes G.J., Saurat J.-H., Siegenthaler G.;
 RT "Characterization of two post-translationally processed forms of
 human serum retinol-binding protein: altered ratios in chronic renal
 failure.";
 RL J. Lipid Res. 36:1247-1253(1995).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=91017498; PubMed=2217163;
 RA Cowan S.W., Newcomer M.E., Jones T.A.;
 RT "Crystallographic refinement of human serum retinol binding protein
 at 2-A resolution.";
 RL Proteins 8:44-61(1990).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY.
 RX MEDLINE=92322903; PubMed=1623143;
 RA Monaco H.L., Zanotti G.;
 RT "Three-dimensional structure and active site of three hydrophobic
 molecule-binding proteins with significant amino acid sequence
 similarity.";
 RL Biopolymers 32:457-465(1992).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS) OF COMPLEX WITH TTR.
 RX MEDLINE=99162254; PubMed=10052934;
 RA Naylor H.M., Newcomer M.E.;
 RT "The structure of human retinol-binding protein (RBP) with its
 carrier protein transthyretin reveals an interaction with the carboxy
 terminus of RBP.";
 RL Biochemistry 38:2647-2653(1999).
 RN [12]
 RP VARIANTS RBP DEFICIENCY ASN-59 AND ASP-93.
 RX MEDLINE=99103495; PubMed=988420;
 RA Seeliger M.W., Biesalski H.K., Wissing B., Gollnick H., Gielen S.,
 RT "Phenotype in retinol deficiency due to a hereditary defect in retinol
 binding protein synthesis.";
 RL Invest. Ophthalmol. Vis. Sci. 40:3-11(1999).
 RN [13]
 RP CHARACTERIZATION OF VARIANTS ASN-59 AND ASP-93.
 RX MEDLINE=99247593; PubMed=10232633;
 RA Biesalski H.K., Frank J., Beck S.C., Heinrich F., Illek B., Reiflen R.,
 RT Gollnick H., Seeliger M.W., Wissing B., Zrenner E.;
 RL "Biochemical but not clinical vitamin A deficiency results from
 mutations in the gene for retinol binding protein.";
 RN Am. J. Clin. Nutr. 69:931-936(1999).
 RP -1- FUNCTION: Delivers retinol from the liver stores to the peripheral
 tissues. In plasma, the RBP-retinol complex interacts with
 transthyretin, this prevents its loss by filtration through the
 kidney glomeruli.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MASS SPECTROMETRY: MW=21063.46; MW_ERR=1.88; METHOD=Electrospray;
 CC RANGE=17-109.
 CC -1- DISEASE: DEFECTS IN RBP4 ARE A CAUSE OF RETINOL-BINDING PROTEIN
 CC DEFICIENCY WHICH CAUSE NIGHT VISION PROBLEMS.
 CC -1- DISEASE: A DEFICIENCY OF VITAMIN A BLOCKS SECRETION OF THE BINDING
 CC PROTEIN POSTTRANSLATIONALLY AND RESULTS IN DEFECTIVE DELIVERY AND
 CC SUPPLY OF VITAMIN TO THE EPIDERMAL CELLS (A CONDITION ASSOCIATED
 CC WITH A DERMATOSIS).

CC -1- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.
 CC -1- DATABASE: NAME=Mutations of the RBP4 gene.
 CC NOTE=Retina International's Scientific Newsletter;
 CC WWW="http://www.retina-international.com/sci-news/rbp4mut.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X00129; CAA24959.1; -;
 DR EMBL: BC020633; AAH20633.1; -;
 DR EMBL: X02775; CAA26353.1; -;
 DR EMBL: X02824; CAA64489.1; -;
 DR EMBL: AF119868; AAF69622.1; ALT_INIT.
 DR EMBL: AF025334; AAC02945.1; -;
 DR EMBL: AF025335; AAC02946.1; -;
 DR PIR: A03223; VAHU
 DR PIR: S06278; S06278.
 DR PIR: A27786; A27786.
 DR PDB: 1BRP; 15-JUL-91.
 DR PDB: 1BRP; 31-JAN-94.
 DR PDB: 1BRQ; 31-JAN-94.
 DR PDB: 1ONB; 09-APR-99.
 DR SWISS-2DPAGE; P02753; HUMAN.
 DR Slena-2DPAGE; P02753; -;
 DR Genew; HGNC:9922; RBP4.
 DR MIM: 180250; -;
 DR InterPro: IPR002345; Lipocalin.
 DR InterPro: IPR000566; Lipocalin_cytfabp.
 DR Pfam: PF00061; Lipocalin; 1.
 DR PRINTS: PR00179; LIPOCALIN.
 DR PROSITE: PS00213; LIPOCALIN; 1.
 KW Plasma; Vitamin A; Retinol-binding; Transport; Liver; Signal;
 KW Lipocalin; Disease mutation; vision; 3D-structure.
 FT SIGNAL 1 18
 FT CHAIN 19 201
 FT DISULFID 22 178
 FT DISULFID 88 192
 FT DISULFID 138 147
 FT VARIANT 59 59
 FT VARIANT 93 93
 FT CONFLICT 8 8
 FT CONFLICT 13 17
 FT HELIX 23 23
 FT HELIX 24 26
 FT TURN 31 32
 FT HELIX 35 38
 FT STRAND 40 48
 FT STRAND 60 66
 FT TURN 67 69
 FT STRAND 70 80
 FT TURN 82 83
 FT STRAND 86 96
 FT TURN 101 102
 FT STRAND 103 110
 FT TURN 113 114
 FT STRAND 117 127
 FT STRAND 132 141
 FT TURN 143 144
 FT STRAND 146 156
 FT TURN 159 160
 FT HELIX 164 176
 FT TURN 177 178
 FT TURN 180 181
 FT STRAND 184 185
 FT SEQUENCE 201 AA; 23044 MW; 6EA8213E863FCDF8 CRC64;
 I -> N (IN RBP DEFICIENCY).
 /FTID-VAR_009276.
 G -> D (IN RBP DEFICIENCY).
 F -> L (IN REF. 1 AND 3).
 LGSGR -> WAA (IN REF. 1 AND 3).

Query Match 4.5%; Score 6; DB 1; Length 201;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 108 AOKIVR 113
Db 166 AOKIVR 171

RESULT 32

YNIC_ECOLI STANDARD; PRT; 222 AA.

AC P77247; P78167;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ynic.
GN YNIC OR B1727.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.

OC NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blotner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[2]

SEQUENCE FROM N.A.

RP STRAIN-K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alda H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashiwato K., Kimura S., Kitagawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
RA Sempel G., Seki Y., Sivasubraman S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horichi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
CC -1- SIMILARITY: BELONGS TO THE CBBY/CBBZ/GPH/YIEH FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AEO00267; AAC74797.1; -;
DR EMBL; D90814; BAA15502.1; -;
DR EMBL; D90815; BAA15508.1; -;
DR EMBL; D90816; BAA15511.1; -;
DR EcoGene; EGI3988; ynic.
DR InterPro; IPR001454; Hlgase/hydrilase.
DR Pfam; PF00702; Hydrolase; 1.
KW Hypothetical protein; Complete proteome.
SO SEQUENCE 222 AA; 24330 MW; 76FE1F2A331476A7 CRC64;

Query Match 4.5%; Score 6; DB 1; Length 222;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 90 STRELT 95
Db 211 STRELT 216

RESULT 33

Y647_HAEIN STANDARD; PRT; 238 AA.

AC Q57424; O05028;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein H10647.
GN H10647.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.

OC NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Ullrichback T.R., Hanna M.C., Nguyen D.T., Saudke D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd.";
RL Science 269:496-512(1995).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE MGTC / SABB FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; U32748; AAC22307.1; -;
DR TIGR; H10647; -;
DR InterPro; IPR003416; MgtC/SapB_tsrprt.
DR Pfam; PF02308; MgtC; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 16 36 POTENTIAL.
FT TRANSMEM 44 64 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 123 143 POTENTIAL.
SO SEQUENCE 238 AA; 25892 MW; 923429C32EC630E8 CRC64;

Query Match 4.5%; Score 6; DB 1; Length 238;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 94 LTTAAI 99
Db 111 LTTAAI 116

RESULT 34

YGIP_YEAST STANDARD; PRT; 239 AA.

AC P53223;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 27.6 kDa protein in RPL26B-ACB1 intergenic region.
GN YGR036C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

```

OX NCB1_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288c;
RX MEDLINE=97435481; PubMed=9290212;
RA Rieger M., Bruckner M., Schaefer M., Mueller-Auer S.;
RT *Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT chromosome VII.
RL Yeast 13:1077-1090(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 272821; CAA97024.1; -.
DR SGD: S0003268; CAX4.
DR InterPro: IPR000326; PA_PTPase.
DR Pfam: PF01569; PAP2; 1.
DR SMART: SM00014; ac1apc; 1.
DR Hypothetical protein; Transmembrane.
KW TRANSMEM 35 55 POTENTIAL.
FT TRANSMEM 132 152 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
SQ SEQUENCE 239 AA; 27649 MW; C8EC49CBFC08F8F8 CRC64;

Query Match
Best Local Similarity 4.5%; Score 6; DB 1; Length 239;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 97 AAINPN 102
DB 6 AAINPN 11

RESULT 35
Y564_METTH STANDARD; PRT; 246 AA.
ID Y564_METTH
AC 026664;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MTH564.
GN MTH564.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
ON NCB1_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Delta H:
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bahrizadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pochler B., Olu D.,
RA Spadafora R., Vlcare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Sefer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- SIMILARITY: STRONG, TO M_JANNASCHT M01676.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000839; AAB85070.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 246 AA; 25575 MW; EA6B930A556BE738 CRC64;

Query Match
Best Local Similarity 4.5%; Score 6; DB 1; Length 246;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 120 MLETVN 125
DB 63 MLETVN 68

RESULT 36
YE91_MYCTU STANDARD; PRT; 252 AA.
ID YE91_MYCTU
AC P71772;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein Rv1491c.
GN Rv1491C OR MT1538 OR MTCV277.13C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
ON NCB1_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultson J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Kouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE UPF0043 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 279701; CAB02041.1; -.
DR EMBL: AE007022; AAK45805.1; -.
DR TIGR: MT1538; -.
DR TubercuList: Rv1491c; -.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 32 52 POTENTIAL.
FT TRANSMEM 64 84 POTENTIAL.

```

FT TRANSMEM 88 108 POTENTIAL.
 FT TRANSMEM 149 169 POTENTIAL.
 FT TRANSMEM 177 197 POTENTIAL.
 FT TRANSMEM 209 229 POTENTIAL.
 SQ SEQUENCE 252 AA: 26575 MW: 3663BE13217B8201 CRC64;

Query Match 4.5%; Score 6; DB 1; Length 252;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 TAFTIA 18
 DB 87 TAFTIA 92

RESULT 37

ID GLTF_ECOLI STANDARD; PRT; 254 AA.
 AC P28721;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protein gltf precursor.
 GN GLTF OR B3214.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.

OX NCBI_TaxId=562;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-K12;
 RX MEDLINE=93078627; PubMed=1447980;

RA Casano I., Flores N., Valle F., Covarrubias A.A., Bolivar F.;

RT "gltF, a member of the gltBDF operon of Escherichia coli, is involved

in nitrogen-regulated gene expression.";

RL Mol. Microbiol. 6:2733-2741(1992).
 [2]

RN SEQUENCE FROM N.A.

RC STRAIN-K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;

RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,

Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

Mau B., Shao Y.;

The complete genome sequence of Escherichia coli K-12.;

Science 277:1453-1474(1997).

RT Science 277:1453-1474(1997).

RL Science 277:1453-1474(1997).

RN Science 277:1453-1474(1997).

RP Science 277:1453-1474(1997).

RC Science 277:1453-1474(1997).

RX Science 277:1453-1474(1997).

RA Science 277:1453-1474(1997).

RT Science 277:1453-1474(1997).

RL Science 277:1453-1474(1997).

RN Science 277:1453-1474(1997).

RP Science 277:1453-1474(1997).

RC Science 277:1453-1474(1997).

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 94 LTTAAI 99
 DB 7 LTTAAI 12

RESULT 38

ID RM24_YEAST STANDARD; PRT; 258 AA.
 AC P36525;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE 60S ribosomal protein L24, mitochondrial precursor (Yml24).
 GN MRPL24 OR YMR193W OR YMG646.05.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxId=4932;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-S288c / AB972;
 RA Pearson D., Bowman S., Barrell B.G., Rajandream M.A.;

RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 [2]

RP SEQUENCE OF 22-49 AND 72-87.

RX MEDLINE=91285106; PubMed=2060626;
 RA Grohmann L., Graack H.-R., Krut V., Choi T., Goldschmidt-Reisin S.,

Kitakawa M.;

*Extended N-terminal sequencing of proteins of the large ribosomal

subunit from yeast mitochondria.";

FEBS Lett. 284:51-56(1991).

RT FEBS Lett. 284:51-56(1991).

RL FEBS Lett. 284:51-56(1991).

RN FEBS Lett. 284:51-56(1991).

RP FEBS Lett. 284:51-56(1991).

RC FEBS Lett. 284:51-56(1991).

RX FEBS Lett. 284:51-56(1991).

RA FEBS Lett. 284:51-56(1991).

RT FEBS Lett. 284:51-56(1991).

RL FEBS Lett. 284:51-56(1991).

RN FEBS Lett. 284:51-56(1991).

RP FEBS Lett. 284:51-56(1991).

RC FEBS Lett. 284:51-56(1991).

RX FEBS Lett. 284:51-56(1991).

RA FEBS Lett. 284:51-56(1991).

RT FEBS Lett. 284:51-56(1991).

RL FEBS Lett. 284:51-56(1991).

RN FEBS Lett. 284:51-56(1991).

RP FEBS Lett. 284:51-56(1991).

RC FEBS Lett. 284:51-56(1991).

RX FEBS Lett. 284:51-56(1991).

RA FEBS Lett. 284:51-56(1991).

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 94 LTTAAI 99
 DB 7 LTTAAI 12

RESULT 39

ID YFAP_ECOLI STANDARD; PRT; 258 AA.
 AC P76462;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein yfap precursor.

RP SEQUENCE FROM N.A.

RC STRAIN-S288c / AB972;
 RA Pearson D., Bowman S., Barrell B.G., Rajandream M.A.;

RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 [2]

RP SEQUENCE OF 22-49 AND 72-87.

RX MEDLINE=91285106; PubMed=2060626;
 RA Grohmann L., Graack H.-R., Krut V., Choi T., Goldschmidt-Reisin S.,

Kitakawa M.;

*Extended N-terminal sequencing of proteins of the large ribosomal

subunit from yeast mitochondria.";

FEBS Lett. 284:51-56(1991).

RT FEBS Lett. 284:51-56(1991).

RL FEBS Lett. 284:51-56(1991).

RN FEBS Lett. 284:51-56(1991).

RP FEBS Lett. 284:51-56(1991).

RC FEBS Lett. 284:51-56(1991).

RX FEBS Lett. 284:51-56(1991).

RA FEBS Lett. 284:51-56(1991).

RT FEBS Lett. 284:51-56(1991).

RL FEBS Lett. 284:51-56(1991).

RN FEBS Lett. 284:51-56(1991).

RP FEBS Lett. 284:51-56(1991).

RC FEBS Lett. 284:51-56(1991).

RX FEBS Lett. 284:51-56(1991).

RA FEBS Lett. 284:51-56(1991).

RT FEBS Lett. 284:51-56(1991).

RL FEBS Lett. 284:51-56(1991).

RN FEBS Lett. 284:51-56(1991).

RP FEBS Lett. 284:51-56(1991).

RC FEBS Lett. 284:51-56(1991).

RX FEBS Lett. 284:51-56(1991).

RA FEBS Lett. 284:51-56(1991).

```

GN VFAP OR B2225.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_Taxid=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RL "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000312; AAC75285.1; -.
DR Ecogene; EGI4078; yfap.
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 19
FT CHAIN 20 258 HYPOTHETICAL PROTEIN VFAP.
SQ SEQUENCE 258 AA; 28303 MW; 911F2B5C433E5F41 CRC64;

Query Match 4.5%; Score 6; DB 1; Length 258;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 92 LETVNA 97
| | | | |
DB 217 LETVNA 222

RESULT 40
RFA4_HUMAN
ID RFA4_HUMAN STANDARD; PRT; 261 AA.
AC Q13156;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Replication protein A 30 kDa subunit (RP-A) (RP-A) (Replication
DE factor-A protein 4).
GN RPA4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE-95280910; PubMed-7760808;
RA Keshav K.F., Chen C., Dutta A.;
RT "Rpa4, a homolog of the 34-kilodalton subunit of the replication
RT protein A complex";
RL Mol. Cell. Biol. 15:3119-3128(1995).
RN [2]
RP REVISION TO 10.
RP Keshav K.F.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ABSOLUTELY REQUIRED FOR SIMIAN VIRUS 40 DNA REPLICATION
CC IN VITRO. IT PARTICIPATES IN A VERY EARLY STEP IN INITIATION. RP-A
CC IS A SINGLE-STRANDED DNA-BINDING PROTEIN (BY SIMILARITY).
CC -1- SUBUNIT: HETEROTRIMER OF 70, 32/30, AND 14 KDA CHAINS. THE DNA-
CC BINDING ACTIVITY MAY RESIDE EXCLUSIVELY ON THE 70 KDA SUBUNIT.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN PLACENTAL AND
CC COLON MUCOSA.

```

```

CC -1- SIMILARITY: TO OTHER SPECIES RFA2/RPA2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U24186; AAB08488.2; -.
DR Interpro: IPR004365; trna_antl.
DR Pfam: PF01336; trna_antl.1.
KW DNA replication; Nuclear protein; Alternative splicing.
SQ SEQUENCE 261 AA; 28868 MW; 6A925FAEDBE21718 CRC64;

Query Match 4.5%; Score 6; DB 1; Length 261;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 121 LETVNA 126
| | | | |
DB 159 LETVNA 164

```

Search completed: April 28, 2003, 16:13:01
 Job time : 29 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 16:10:58 ; Search time 85 Seconds
(without alignments)
324.827 Million cell updates/sec

Title: US-10-048-196-2

Perfect score: 134
Sequence: 1 MMKILYVATLMTAFTLASC.....SLKPCMLETVNAFIVPTTTR 134

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

- 1: SP archaea:*
- 2: SP bacteria:*
- 3: SP fungi:*
- 4: SP human:*
- 5: SP invertebrate:*
- 6: SP mammal:*
- 7: SP mhc:*
- 8: SP organelle:*
- 9: SP phage:*
- 10: SP plant:*
- 11: SP rodent:*
- 12: SP virus:*
- 13: SP vertebrate:*
- 14: SP unclassified:*
- 15: SP yvirus:*
- 16: SP bacteriaph:*
- 17: SP archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	6.0	100	16	08YHV9
2	8	6.0	120	16	092RM9
3	8	6.0	330	4	014882
4	8	6.0	543	4	014879
5	8	6.0	622	4	014881
6	8	6.0	716	4	09NYE4
7	8	6.0	864	12	08U2D5
8	8	6.0	3570	4	099552
9	8	5.2	64	16	097FE31
10	8	5.2	132	16	08YFH5
11	8	5.2	174	16	08YIT4
12	8	5.2	175	16	09KKI0
13	8	5.2	185	17	0975X8
14	8	5.2	186	16	08YGD7
15	8	5.2	213	10	09SAG6
16	8	5.2	215	5	09V3Y0

17	7	5.2	222	5	08P5S3	08T5S3 trypanosoma
18	7	5.2	233	15	09Q348	09Q348 human immun
19	7	5.2	246	13	090877	090877 gallus gall
20	7	5.2	246	15	085408	085408 avian endog
21	7	5.2	255	2	09XDX8	09XDX8 rhodospo
22	7	5.2	277	5	016217	016217 caenorhabd
23	7	5.2	302	16	09K059	09K059 vibrio chol
24	7	5.2	320	5	093892	093892 caenorhabd
25	7	5.2	362	2	09RCH0	09RCH0 vibrio chol
26	7	5.2	389	3	096X43	096X43 kluyveromy
27	7	5.2	399	5	09GSG9	09GSG9 dictyostel
28	7	5.2	436	16	08U7N5	08U7N5 agrobacteri
29	7	5.2	459	8	09X111	09X111 falco pereg
30	7	5.2	499	17	026204	026204 methanobact
31	7	5.2	501	17	08T0Y6	08T0Y6 methanosarc
32	7	5.2	577	15	007453	007453 rous sarcom
33	7	5.2	595	15	003819	003819 rous sarcom
34	7	5.2	600	15	083129	083129 rous sarcom
35	7	5.2	601	15	090VD8	090VD8 avian myelo
36	7	5.2	601	15	090VD7	090VD7 avian leuko
37	7	5.2	601	15	090VD6	090VD6 avian leuko
38	7	5.2	601	15	090VD5	090VD5 avian leuko
39	7	5.2	601	15	064984	064984 rous sarcom
40	7	5.2	604	15	083132	083132 avian myelo
41	7	5.2	605	17	08TR19	08TR19 methanosarc
42	7	5.2	607	15	083134	083134 avian myelo
43	7	5.2	611	15	098WV7	098WV7 avian leuko
44	7	5.2	613	15	092955	092955 rous sarcom
45	7	5.2	614	15	098WV9	098WV9 avian leuko
46	7	5.2	616	15	098WV1	098WV1 avian leuko
47	7	5.2	618	10	09LV74	09LV74 arabidopsis
48	7	5.2	620	10	09ASS4	09ASS4 arabidopsis
49	7	5.2	620	10	08W4H2	08W4H2 arabidopsis
50	7	5.2	633	5	095P18	095P18 aedes aegy
51	7	5.2	633	5	08WOL6	08WOL6 aedes aegy
52	7	5.2	633	5	08WOL5	08WOL5 aedes aegy
53	7	5.2	633	5	08WOL4	08WOL4 aedes aegy
54	7	5.2	633	5	08WOL3	08WOL3 aedes aegy
55	7	5.2	633	5	08WOL2	08WOL2 aedes aegy
56	7	5.2	633	5	016894	016894 aedes aegy
57	7	5.2	725	15	08XT23	08XT23 ralslonia s
58	7	5.2	728	4	060561	060561 homo sapien
59	7	5.2	906	13	08UVU1	08UVU1 xenopus lae
60	7	5.2	928	16	08Y604	08Y604 listeria mo
61	7	5.2	988	13	08UVU2	08UVU2 xenopus lae
62	7	5.2	1039	17	097V47	097V47 sulfobus
63	7	5.2	1077	5	09NTG7	09NTG7 drosophila
64	7	5.2	1848	5	09YU86	09YU86 drosophila
65	7	5.2	2117	3	096U11	096U11 neurospora
66	7	5.2	3810	5	09U943	09U943 locusta mlg
67	7	5.2	4001	5	09N2P7	09N2P7 drosophila
68	7	5.2	4796	5	09NLAB	09NLAB drosophila
69	7	5.2	4796	5	09W055	09W055 drosophila
70	7	5.2	16215	5	09NES3	09NES3 drosophila
71	7	4.5	27	5	096320	096320 tryptesea 1
72	7	4.5	29	2	09R4B7	09R4B7 streptomyc
73	7	4.5	37	10	09ZP81	09ZP81 raphanus sa
74	7	4.5	44	10	09S928	09S928 alnus glutl
75	7	4.5	53	10	08RZG7	08RZG7 oryza sativ
76	7	4.5	64	16	0932D0	0932D0 staphylococ
77	7	4.5	65	10	09F0M4	09F0M4 brassica na
78	7	4.5	72	1	09Y8K6	09Y8K6 methanosarc
79	7	4.5	84	17	096Z41	096Z41 sulfobus
80	7	4.5	86	16	08X3A0	08X3A0 escherichia
81	7	4.5	92	15	071536	071536 human immun
82	7	4.5	92	15	071537	071537 human immun
83	7	4.5	92	15	071538	071538 human immun
84	7	4.5	92	15	071539	071539 human immun
85	7	4.5	92	15	071540	071540 human immun
86	7	4.5	92	15	071541	071541 human immun
87	7	4.5	92	15	071542	071542 human immun
88	7	4.5	92	15	071543	071543 human immun
89	7	4.5	92	15	071544	071544 human immun

90	6	4.5	92	15	071545	071545 human immun	163	6	4.5	172	16	084854	084854 chlamydia t
91	6	4.5	92	15	071546	071546 human immun	164	6	4.5	172	16	09PL73	09PL73 chlamydia m
92	6	4.5	92	15	071547	071547 human immun	165	6	4.5	173	2	053656	053656 stigmella
93	6	4.5	92	15	071548	071548 human immun	166	6	4.5	177	10	094085	094085 aradidopsis
94	6	4.5	92	15	071549	071549 human immun	167	6	4.5	179	16	097DL8	097DL8 clostridium
95	6	4.5	92	15	071550	071550 human immun	168	6	4.5	181	17	08TRP1	08TRP1 methanosarc
96	6	4.5	92	15	071551	071551 human immun	169	6	4.5	182	4	015402	015402 homo sapien
97	6	4.5	92	15	071626	071626 human immun	170	6	4.5	183	2	09RRS3	09RRS3 vibrio algi
98	6	4.5	92	15	071627	071627 human immun	171	6	4.5	183	10	09LSM4	09LSM4 aradidopsis
99	6	4.5	93	3	09P718	09P718 schizosacch	172	6	4.5	183	16	09CLG4	09CLG4 pasteurella
100	6	4.5	95	2	093EJ0	093EJ0 helicobacte	173	6	4.5	184	5	09VV70	09VV70 drosophila
101	6	4.5	95	10	08MB95	08MB95 brassica na	174	6	4.5	184	12	056972	056972 newcastle d
102	6	4.5	96	13	09DFE6	09DFE6 oncorhynch	175	6	4.5	184	12	056973	056973 newcastle d
103	6	4.5	96	16	08ZMH5	08ZMH5 salmonella	176	6	4.5	184	12	056974	056974 newcastle d
104	6	4.5	96	16	08Z488	08Z488 salmonella	177	6	4.5	184	12	056975	056975 newcastle d
105	6	4.5	99	16	09K218	09K218 chlamydia p	178	6	4.5	184	12	056976	056976 newcastle d
106	6	4.5	102	5	09U3J9	09U3J9 caenorhabd	179	6	4.5	184	12	056977	056977 newcastle d
107	6	4.5	103	5	09U3J9	09U3J9 caenorhabd	180	6	4.5	184	12	056980	056980 newcastle d
108	6	4.5	107	16	09CJU4	09CJU4 caenorhabd	181	6	4.5	184	12	056981	056981 newcastle d
109	6	4.5	108	5	P91138	09G08 homo sapien	182	6	4.5	184	12	056982	056982 newcastle d
110	6	4.5	109	4	09NQ08	09G610 homo sapien	183	6	4.5	190	2	08PFB9	08PFB9 escherichia
111	6	4.5	110	17	058149	09G610 homo sapien	184	6	4.5	197	4	09UDB9	09UDB9 escherichia
112	6	4.5	117	16	005574	058149 mycobacteri	185	6	4.5	197	4	09UDB9	09UDB9 homo sapien
113	6	4.5	118	17	097V14	09V14 sulfolobus	186	6	4.5	198	1	0977U1	0977U1 sulfolobus
114	6	4.5	119	5	020208	020208 caenorhabd	187	6	4.5	198	16	09A4X7	09A4X7 caulobacter
115	6	4.5	124	13	09W618	09W618 xiphophorus	188	6	4.5	205	16	09P4X7	09P4X7 leishmania
116	6	4.5	124	13	P70067	P70067 xiphophorus	189	6	4.5	206	16	09RMW2	09RMW2 delnoccocus
117	6	4.5	126	8	09TKN7	09TKN7 exacum affi	190	6	4.5	206	16	09P4X7	09P4X7 rhizobium m
118	6	4.5	129	5	09NMJ7	09NMJ7 onchocerca	191	6	4.5	206	16	09P4X7	09P4X7 xylella fas
119	6	4.5	130	15	088013	088013 chimpazee	192	6	4.5	206	16	09P4X7	09P4X7 caenorhabd
120	6	4.5	131	15	090EX7	090EX7 simlan immu	193	6	4.5	211	8	033764	033764 allomyces m
121	6	4.5	131	15	090EX7	090EX7 simlan immu	194	6	4.5	212	10	0941M1	0941M1 oryza sativ
122	6	4.5	132	4	09H6P2	09H6P2 homo sapien	195	6	4.5	214	16	08YNN5	08YNN5 anabena sp
123	6	4.5	132	4	09H6B10	09H6B10 homo sapien	196	6	4.5	215	16	08YNN5	08YNN5 streptomyce
124	6	4.5	134	16	08Z588	08Z588 salmonella	197	6	4.5	215	16	09K4E2	09K4E2 streptomyce
125	6	4.5	136	17	08TNR4	08TNR4 methanosarc	198	6	4.5	216	12	091CW7	091CW7 viral hemor
126	6	4.5	139	16	092TK3	092TK3 rhizobium m	199	6	4.5	219	2	09KJ15	09KJ15 viral hemor
127	6	4.5	139	16	08YFA9	08YFA9 bruceella me	200	6	4.5	220	12	091CW9	091CW9 viral hemor
128	6	4.5	140	11	09W026	09W026 cavia porce	201	6	4.5	222	12	082124	082124 viral hemor
129	6	4.5	141	2	0939J1	0939J1 vibrio chol	202	6	4.5	222	12	082125	082125 viral hemor
130	6	4.5	142	16	08YEP1	08YEP1 bruceella me	203	6	4.5	222	12	090B88	090B88 viral hemor
131	6	4.5	144	2	082910	082910 escherichia	204	6	4.5	222	12	09WAB8	09WAB8 viral hemor
132	6	4.5	144	2	09RGZ4	09RGZ4 bacillus ps	205	6	4.5	222	12	090B88	090B88 viral hemor
133	6	4.5	145	2	09ZGR7	09ZGR7 escherichia	206	6	4.5	222	12	09PWJ3	09PWJ3 viral hemor
134	6	4.5	146	15	085706	085706 rous-associ	207	6	4.5	222	12	088475	088475 viral hemor
135	6	4.5	146	16	08YTD4	08YTD4 anabena sp	208	6	4.5	222	16	08Z6H5	08Z6H5 viral hemor
136	6	4.5	146	16	044659	044659 bacillus ha	209	6	4.5	222	16	08Z6H5	08Z6H5 salmonella
137	6	4.5	147	12	098175	098175 moluscum c	210	6	4.5	222	16	08YNN6	08YNN6 anabena sp
138	6	4.5	148	2	09AKC7	09AKC7 rickettsia	211	6	4.5	222	16	08YNN6	08YNN6 anabena sp
139	6	4.5	148	2	09AKH4	09AKH4 rickettsia	212	6	4.5	222	16	08YNN6	08YNN6 anabena sp
140	6	4.5	148	2	09AKM9	09AKM9 rickettsia	213	6	4.5	222	16	09RDS7	09RDS7 streptomyce
141	6	4.5	148	10	09C527	09C527 aradidopsis	214	6	4.5	223	12	09WM16	09WM16 newcastle d
142	6	4.5	148	16	09ZDM9	09ZDM9 rickettsia	215	6	4.5	223	12	09WMT3	09WMT3 newcastle d
143	6	4.5	148	16	09ZJ11	09ZJ11 rickettsia	216	6	4.5	223	16	09CPC3	09CPC3 pasteurella
144	6	4.5	149	2	09X549	09X549 corynebacte	217	6	4.5	223	16	08YPD3	08YPD3 anabena sp
145	6	4.5	153	5	09U986	09U986 drosophila	218	6	4.5	226	12	098332	098332 microplitis
146	6	4.5	153	13	090YU4	090YU4 ictalurus p	219	6	4.5	227	10	09ZPS8	09ZPS8 aradidopsis
147	6	4.5	153	13	090YU6	090YU6 staphylococ	220	6	4.5	233	16	08Y034	08Y034 ralsconia s
148	6	4.5	154	16	08R7R8	08R7R8 thermomane	221	6	4.5	234	13	0919W1	0919W1 dendragapus
149	6	4.5	155	2	033696	033696 streptococ	222	6	4.5	234	13	0919W1	0919W1 dendragapus
150	6	4.5	158	10	094H22	094H22 oryza sativ	223	6	4.5	235	16	099Z55	099Z55 streptococ
151	6	4.5	159	16	066447	066447 aquifex aeo	224	6	4.5	235	16	099Z55	099Z55 streptococ
152	6	4.5	160	10	039425	039425 betula verr	225	6	4.5	236	5	0965N0	0965N0 caenorhabd
153	6	4.5	161	17	09YD02	09YD02 aeropyrum p	226	6	4.5	236	12	067552	067552 garlic late
154	6	4.5	163	16	08YXT8	08YXT8 anabaena sp	227	6	4.5	239	13	0919W5	0919W5 dendragapus
155	6	4.5	163	16	08YXT8	08YXT8 anabaena sp	228	6	4.5	239	13	0919W5	0919W5 dendragapus
156	6	4.5	165	10	09AYE8	09AYE8 brassica ca	229	6	4.5	240	16	08YJH1	08YJH1 pasteurella
157	6	4.5	165	16	08R7R0	08R7R0 thermomane	230	6	4.5	241	3	096183	096183 homo sapien
158	6	4.5	167	10	09LH78	09LH78 aradidopsis	231	6	4.5	241	3	0919S5	0919S5 phasianus c
159	6	4.5	170	10	08SAV0	08SAV0 oryza sativ	232	6	4.5	243	13	0919Z0	0919Z0 bonasa umbe
160	6	4.5	171	10	096LH4	096LH4 sulfolobus	233	6	4.5	243	13	0919Z0	0919Z0 centrocercu
161	6	4.5	171	10	09LH11	09LH11 aradidopsis	234	6	4.5	245	10	040594	040594 nicotiana t
162	6	4.5	171	16	09PRR4	09PRR4 campylobact	235	6	4.5	245	10	040594	040594 nicotiana t

235	6	4.5	245	16	09cjp3	pasteurella	309	6	4.5	286	16	092WR4	092WR4	rhizobium m
237	6	4.5	246	10	094JN9	094JN9	310	6	4.5	287	17	09CGC8	09CGC8	lactococcus
238	6	4.5	246	13	091872	091872	311	6	4.5	287	17	09VIM8	09VIM8	pyrococcus
239	6	4.5	246	13	0919T0	0919T0	312	6	4.5	288	2	09K581	09K581	klebsiella
240	6	4.5	246	13	0919S8	0919S8	313	6	4.5	289	16	09ZC43	09ZC43	yersinia pe
241	6	4.5	246	13	0919S7	0919S7	314	6	4.5	291	16	P6851	P6851	mycobacteri
242	6	4.5	246	13	0919S6	0919S6	315	6	4.5	292	2	09KMO6	09KMO6	rhodococcus
243	6	4.5	247	17	08T101	08T101	316	6	4.5	292	16	09X7Q1	09X7Q1	streptomyces
244	6	4.5	248	16	08XW2	08XW2	317	6	4.5	293	10	094H55	094H55	oryza sativ
245	6	4.5	249	16	08YS76	08YS76	318	6	4.5	293	16	08XAR4	08XAR4	escherichia
246	6	4.5	249	16	09RK49	09RK49	319	6	4.5	294	5	018893	018893	caenorhabd
247	6	4.5	251	4	09H7V3	09H7V3	320	6	4.5	294	5	09V504	09V504	drosophila
248	6	4.5	251	13	0919X5	0919X5	321	6	4.5	294	13	09V195	09V195	bonasa umbe
249	6	4.5	251	13	0919W2	0919W2	322	6	4.5	295	10	022158	022158	arabidopsis
250	6	4.5	252	3	094030	094030	323	6	4.5	295	16	09A5E6	09A5E6	caulobacter
251	6	4.5	253	16	09EWY3	09EWY3	324	6	4.5	297	16	08U9P0	08U9P0	agrobacteri
252	6	4.5	254	12	09WMT7	09WMT7	325	6	4.5	298	16	09KJS2	09KJS2	streptomyces
253	6	4.5	254	12	09WJR1	09WJR1	326	6	4.5	299	2	09R780	09R780	lactococcus
254	6	4.5	254	12	09WMT1	09WMT1	327	6	4.5	299	4	09NMY4	09NMY4	homo sapien
255	6	4.5	254	16	09RYZ3	09RYZ3	328	6	4.5	299	5	09V8N2	09V8N2	drosophila
256	6	4.5	255	9	09MBV4	09MBV4	329	6	4.5	299	16	08ZPK8	08ZPK8	salmonella
257	6	4.5	255	16	09KO14	09KO14	330	6	4.5	299	16	08Z6X7	08Z6X7	salmonella
258	6	4.5	255	16	09GCS4	09GCS4	331	6	4.5	301	12	091168	091168	newcastle d
259	6	4.5	256	12	09WMT4	09WMT4	332	6	4.5	303	5	09V8N3	09V8N3	drosophila
260	6	4.5	256	12	09YN77	09YN77	333	6	4.5	304	12	09IML7	09IML7	choristoneu
261	6	4.5	256	12	09WMT2	09WMT2	334	6	4.5	304	16	08Y8N3	08Y8N3	listeria mo
262	6	4.5	258	12	09WBJ2	09WBJ2	335	6	4.5	305	16	08X1W8	08X1W8	clostridium
263	6	4.5	258	16	08XE38	08XE38	336	6	4.5	306	16	09RY72	09RY72	deinococcus
264	6	4.5	260	9	09XJ66	09XJ66	337	6	4.5	307	16	09ZLK2	09ZLK2	rhizobium m
265	6	4.5	260	9	09BOC2	09BOC2	338	6	4.5	308	2	032718	032718	klebsiella
266	6	4.5	260	9	09BOC1	09BOC1	339	6	4.5	311	5	09G932	09G932	tetrahymena
267	6	4.5	260	9	09BO07	09BO07	340	6	4.5	311	9	038139	038139	bacterioph
268	6	4.5	261	16	09A941	09A941	341	6	4.5	311	9	048490	048490	bacterioph
269	6	4.5	261	2	04G988	04G988	342	6	4.5	311	17	08TPC8	08TPC8	methanosarc
270	6	4.5	261	4	09BQND	09BQND	343	6	4.5	312	2	08VW71	08VW71	rhodospseudo
271	6	4.5	261	12	091HX2	091HX2	344	6	4.5	313	16	09PMNO	09PMNO	campylobact
272	6	4.5	261	12	091HS3	091HS3	345	6	4.5	314	10	08VXQ9	08VXQ9	chlorella f
273	6	4.5	261	16	08YZI3	08YZI3	346	6	4.5	314	17	08TNGO	08TNGO	methanosarc
274	6	4.5	261	16	054146	054146	347	6	4.5	314	17	08TLB2	08TLB2	methanosarc
275	6	4.5	261	17	028627	028627	348	6	4.5	315	11	09D4W2	09D4W2	mus musculu
276	6	4.5	261	17	09HOK8	09HOK8	349	6	4.5	316	16	08XLR4	08XLR4	clostridium
277	6	4.5	262	17	08T246	08T246	350	6	4.5	317	3	09GUL16	09GUL16	neurospora
278	6	4.5	263	12	09WMI5	09WMI5	351	6	4.5	318	2	09F8F8	09F8F8	uncultured
279	6	4.5	263	12	09WMI0	09WMI0	352	6	4.5	319	5	062451	062451	caenorhabd
280	6	4.5	267	2	09L7R1	09L7R1	353	6	4.5	320	16	052908	052908	rhizobium m
281	6	4.5	267	16	09A408	09A408	354	6	4.5	322	2	09F7P8	09F7P8	uncultured
282	6	4.5	267	16	08ZKT6	08ZKT6	355	6	4.5	322	2	08RSO6	08RSO6	mannheimia
283	6	4.5	267	16	08Z2T7	08Z2T7	356	6	4.5	323	2	099Q25	099Q25	enterobacte
284	6	4.5	268	16	08YLI0	08YLI0	357	6	4.5	323	16	08YXF8	08YXF8	anabaena sp
285	6	4.5	269	16	08XBD2	08XBD2	358	6	4.5	324	4	09UPK9	09UPK9	homo sapien
286	6	4.5	270	16	09KY29	09KY29	359	6	4.5	325	2	091JP2	091JP2	uncultured
287	6	4.5	270	17	028514	028514	360	6	4.5	325	2	09EXF8	09EXF8	bacterioides
288	6	4.5	271	17	08T847	08T847	361	6	4.5	326	2	070076	070076	acidiphiliu
289	6	4.5	272	12	036276	036276	362	6	4.5	326	4	09UPK8	09UPK8	homo sapien
290	6	4.5	272	12	036289	036289	363	6	4.5	326	10	09XFR3	09XFR3	arabidopsis
291	6	4.5	273	11	0920V5	0920V5	364	6	4.5	326	16	098N25	098N25	rhizobium l
292	6	4.5	273	16	091OK5	091OK5	365	6	4.5	327	5	095CJ1	095CJ1	caenorhabd
293	6	4.5	273	16	08UE93	08UE93	366	6	4.5	328	13	09PRJ2	09PRJ2	brachydanio
294	6	4.5	274	6	09BG12	09BG12	367	6	4.5	329	10	022498	022498	glycine max
295	6	4.5	274	16	08ZNI4	08ZNI4	368	6	4.5	331	10	039954	039954	helianthus
296	6	4.5	275	2	08VUD9	08VUD9	369	6	4.5	332	5	09F1V1	09F1V1	kitasatospo
297	6	4.5	275	11	09CMI0	09CMI0	370	6	4.5	332	16	09GRU9	09GRU9	caenorhabd
298	6	4.5	277	13	0919Y9	0919Y9	371	6	4.5	332	16	08RRY5	08RRY5	deinococcus
299	6	4.5	277	13	0919Y7	0919Y7	372	6	4.5	332	16	098FA3	098FA3	rhizobium l
300	6	4.5	277	13	0919Y6	0919Y6	373	6	4.5	332	16	08XBM6	08XBM6	escherichia
301	6	4.5	277	17	09HKV3	09HKV3	374	6	4.5	333	4	093J08	093J08	streptomyces
302	6	4.5	277	17	09HB76	09HB76	375	6	4.5	333	4	075238	075238	homo sapien
303	6	4.5	280	16	09RKO4	09RKO4	376	6	4.5	334	16	08Y5Z5	08Y5Z5	anabaena sp
304	6	4.5	281	16	09CF98	09CF98	377	6	4.5	335	2	093M61	093M61	staphylococ
305	6	4.5	282	10	09SUI4	09SUI4	378	6	4.5	335	16	075237	075237	homo sapien
306	6	4.5	282	16	08XYG6	08XYG6	379	6	4.5	335	16	09CNM6	09CNM6	pasteurella
307	6	4.5	283	4	000573	000573	380	6	4.5	336	16	09K142	09K142	neisseria m
308	6	4.5	284	16	093J38	093J38	381	6	4.5	336	16			

382	6	4.5	336	16	Q9JST8	Q9JST8 neisseria m	455	5	4.5	372	10	Q9SYT8	Q9SYT8 arabidopsis
383	6	4.5	336	16	Q9Z5C5	Q9Z5C5 staphylococ	456	6	4.5	374	2	Q9SQ24	Q9SQ24 salmonella
384	6	4.5	336	17	Q9J7W0	Q9J7W0 sulfolobus	457	6	4.5	375	2	Q9REY4	Q9REY4 staphylococ
385	6	4.5	337	5	Q76829	Q76829 caenorhabdi	458	6	4.5	375	2	Q9AHX9	Q9AHX9 pseudomonas
386	6	4.5	337	16	Q9ZP25	Q9ZP25 rhizobium m	459	6	4.5	375	5	Q95UN3	Q95UN3 strylongichia
387	6	4.5	338	12	Q99B78	Q99B78 newcastle d	460	6	4.5	375	16	Q9J0T2	Q9J0T2 pseudomonas
388	6	4.5	339	16	Q9XWNT	Q9XWNT thermotoga	461	6	4.5	375	16	Q99SE7	Q99SE7 ralsstonia s
389	6	4.5	339	2	Q9R6W2	Q9R6W2 synecchococ	462	6	4.5	376	16	Q99SE7	Q99SE7 staphylococ
390	6	4.5	340	16	Q9A0J4	Q9A0J4 escherichia	463	6	4.5	377	16	Q8U9V0	Q8U9V0 agrobacteri
391	6	4.5	341	16	Q8XCJ9	Q8XCJ9 yersinia pe	464	6	4.5	377	16	Q8U9V0	Q8U9V0 agrobacteri
392	6	4.5	342	16	Q8ZEC0	Q8ZEC0 agrobacteri	465	6	4.5	377	16	Q8X529	Q8X529 ralsstonia s
393	6	4.5	343	16	Q8U890	Q8U890 candida alb	466	6	4.5	378	16	Q9CM22	Q9CM22 pasteurella
394	6	4.5	344	3	Q96WLC	Q96WLC eremias gra	467	6	4.5	378	16	Q8ZGN7	Q8ZGN7 yersinia pe
395	6	4.5	344	8	P92631	P92631 eremias gra	468	6	4.5	379	3	Q8X0F4	Q8X0F4 neurospora
396	6	4.5	345	8	Q9TC78	Q9TC78 saureia ag	469	6	4.5	379	10	Q9M2R7	Q9M2R7 arabidopsis
397	6	4.5	345	8	Q9TC75	Q9TC75 wetmorena h	470	6	4.5	379	10	Q8S3E2	Q8S3E2 arabidopsis
398	6	4.5	345	8	Q9G497	Q9G497 anolis reco	471	6	4.5	380	2	Q55245	Q55245 synecchococ
399	6	4.5	345	8	Q9G496	Q9G496 anolis reco	472	6	4.5	380	4	Q96882	Q96882 homo sapien
400	6	4.5	345	8	Q8S1J2	Q8S1J2 leptosomus	473	6	4.5	380	4	Q96882	Q96882 homo sapien
401	6	4.5	346	8	Q950B6	Q950B6 empidonax h	474	6	4.5	380	4	Q96103	Q96103 mus musculu
402	6	4.5	346	8	Q950B5	Q950B5 empidonax m	475	6	4.5	380	11	Q9D7P5	Q9D7P5 mus musculu
403	6	4.5	346	8	Q8WB01	Q8WB01 empidonax f	476	6	4.5	380	11	Q921D7	Q921D7 mus musculu
404	6	4.5	346	8	Q8WB00	Q8WB00 empidonax f	477	6	4.5	380	11	Q916W7	Q916W7 mus musculu
405	6	4.5	346	8	Q8WB00	Q8WB00 empidonax f	478	6	4.5	380	16	Q8U9A0	Q8U9A0 mus musculu
406	6	4.5	346	8	Q8WB00	Q8WB00 empidonax f	479	6	4.5	380	16	Q8U9A0	Q8U9A0 mus musculu
407	6	4.5	346	8	Q8WB00	Q8WB00 empidonax f	480	6	4.5	380	16	Q8U9A0	Q8U9A0 mus musculu
408	6	4.5	346	8	Q8WB00	Q8WB00 empidonax f	481	6	4.5	381	13	Q919W3	Q919W3 dendragapus
409	6	4.5	346	8	Q8WB00	Q8WB00 empidonax f	482	6	4.5	381	13	Q919W3	Q919W3 dendragapus
410	6	4.5	346	8	Q8WB00	Q8WB00 empidonax f	483	6	4.5	382	2	Q48713	Q48713 arabidopsis
411	6	4.5	346	8	Q8WB00	Q8WB00 empidonax f	484	6	4.5	386	5	Q9BRP3	Q48713 lactococcus
412	6	4.5	346	8	Q8WB00	Q8WB00 empidonax f	485	6	4.5	386	16	Q8U169	Q9BKP3 caenorhabdi
413	6	4.5	346	8	Q8WB00	Q8WB00 empidonax f	486	6	4.5	387	2	Q93N00	Q8U169 agrobacteri
414	6	4.5	346	8	Q8WB00	Q8WB00 empidonax f	487	6	4.5	387	10	Q93X63	Q93KP3 escherichia
415	6	4.5	346	8	Q8WB00	Q8WB00 empidonax f	488	6	4.5	389	10	Q39688	Q93KP3 brassica na
416	6	4.5	346	8	Q8WB00	Q8WB00 empidonax f	489	6	4.5	390	2	Q9EXA3	Q93688 daucus caro
417	6	4.5	346	8	Q8WB00	Q8WB00 empidonax f	490	6	4.5	390	2	Q9EXA2	Q9EXA3 ochrobactru
418	6	4.5	346	8	Q8WB00	Q8WB00 empidonax f	491	6	4.5	390	2	Q9EXA1	Q9EXA2 ochrobactru
419	6	4.5	346	8	Q8WB00	Q8WB00 empidonax f	492	6	4.5	390	2	Q9EXA0	Q9EXA1 ochrobactru
420	6	4.5	346	8	Q8WB00	Q8WB00 empidonax f	493	6	4.5	390	2	Q9EX99	Q9EXA0 ochrobactru
421	6	4.5	347	2	Q93RP1	Q93RP1 xenorhabdus	494	6	4.5	390	2	Q9EX97	Q9EX99 ochrobactru
422	6	4.5	347	10	Q9Z0N1	Q9Z0N1 arabidopsis	495	6	4.5	390	13	Q919Z4	Q9EX97 ochrobactru
423	6	4.5	348	8	Q9TD50	Q9TD50 cynolebias	496	6	4.5	394	10	Q9XG86	Q919Z4 bonasa bona
424	6	4.5	348	8	Q94M56	Q94M56 gnatholepis	497	6	4.5	395	16	Q9ZEA7	Q9XG86 phleum prat
425	6	4.5	348	8	Q94M45	Q94M45 gnatholepis	498	6	4.5	395	16	Q8RAK7	Q9ZEA7 rickettsia
426	6	4.5	350	2	Q8V8B0	Q8V8B0 uncultured	499	6	4.5	397	16	Q8X2H2	Q8RAK7 thermomanc
427	6	4.5	350	5	Q9UAY5	Q9UAY5 caenorhabdi	500	6	4.5	398	9	Q8SDP1	Q8X2H2 escherichia
428	6	4.5	350	10	Q81850	Q81850 arabidopsis	501	6	4.5	398	13	Q91901	Q8SDP1 pseudomonas
429	6	4.5	352	4	Q15403	Q15403 homo sapien	502	6	4.5	400	4	Q9C0D9	Q91901 lagopus leu
430	6	4.5	352	4	Q08266	Q08266 homo sapien	503	6	4.5	400	16	Q9RK17	Q9C0D9 homo sapien
431	6	4.5	353	16	Q8VAX8	Q8VAX8 bruceella me	504	6	4.5	406	10	Q9XGA0	Q9RK17 bacillus ha
432	6	4.5	355	13	Q9W615	Q9W615 latimeria c	505	6	4.5	407	2	Q93P32	Q9XGA0 spinacla ol
433	6	4.5	358	2	Q9L795	Q9L795 klebsiella	506	6	4.5	409	10	Q9YBBA	Q93P32 acinetobact
434	6	4.5	359	5	Q95W55	Q95W55 pristina le	507	6	4.5	409	10	Q8S2E5	Q9YBBA drosophila
435	6	4.5	360	2	Q9F9U6	Q9F9U6 azoarcus ev	508	6	4.5	409	16	Q8RUP8	Q8S2E5 oryza sativ
436	6	4.5	360	2	Q9FA57	Q9FA57 azoarcus ev	509	6	4.5	409	16	P73556	Q8RUP8 oryza sativ
437	6	4.5	361	5	Q24397	Q24397 drosophila	510	6	4.5	409	16	Q8XVJ7	P73556 synecchocyst
438	6	4.5	362	11	Q9CWM4	Q9CWM4 mus musculu	511	6	4.5	410	5	Q9VG14	Q8XVJ7 ralsstonia s
439	6	4.5	363	12	Q9QB28	Q9QB28 human adeno	512	6	4.5	410	17	Q9XJ85	Q9VG14 drosophila
440	6	4.5	364	2	Q68256	Q68256 leprospira	513	6	4.5	411	10	Q93Z66	Q9XJ85 thermoplasm
441	6	4.5	364	16	Q55548	Q55548 synecchocyst	514	6	4.5	412	16	Q9JTS0	Q93Z66 arabidopsis
442	6	4.5	365	16	Q9A475	Q9A475 caulobacter	515	6	4.5	414	5	Q9B1S6	Q9JTS0 neisseria m
443	6	4.5	365	16	Q8YMA6	Q8YMA6 anabaena sp	516	6	4.5	414	10	Q93YH6	Q9B1S6 actinia sp.
444	6	4.5	366	2	Q937X8	Q937X8 edwardsiell	517	6	4.5	414	10	Q93717	Q93YH6 galdieria s
445	6	4.5	366	2	Q8VP07	Q8VP07 burkholderi	518	6	4.5	414	17	Q9V0F3	Q93717 chondrus cr
446	6	4.5	366	10	Q918M6	Q918M6 brassica na	519	6	4.5	416	16	Q9PM50	Q9V0F3 pyrococcus
447	6	4.5	366	13	Q918M6	Q918M6 brassica na	520	6	4.5	416	16	Q9KSS2	Q9PM50 campylobact
448	6	4.5	367	16	Q82LR2	Q82LR2 dendragapus	521	6	4.5	416	16	Q8X775	Q9KSS2 vibrio chol
449	6	4.5	367	17	Q90ZB3	Q90ZB3 salmonella	522	6	4.5	417	4	Q75236	Q8X775 escherichia
450	6	4.5	368	10	Q48885	Q48885 brassica na	523	6	4.5	419	2	Q90PP3	Q75236 homo sapien
451	6	4.5	368	10	Q8VXW6	Q8VXW6 chara vulga	524	6	4.5	419	4	Q15241	Q90PP3 neisseria m
452	6	4.5	368	17	Q9UZYV	Q9UZYV pyrococcus	525	6	4.5	419	10	Q07795	Q15241 homo sapien
453	6	4.5	370	4	Q9H1N4	Q9H1N4 homo sapien	526	6	4.5	419	13	Q8OHL1	Q07795 sorghum bic
454	6	4.5	370	16	Q97KU0	Q97KU0 clostridium	527	6	4.5	419	16	Q986E2	Q8OHL1 brachydanio
													Q986E2 rhizobium l

528	6	4.5	420	4	09XKJ2	09XKJ2 homo sapien	601	6	4.5	473	10	09LFC2	09LFC2 arabidopsis
529	6	4.5	420	4	09H6P5	09H6P5 homo sapien	602	6	4.5	474	4	08TAK2	08TAK2 homo sapien
530	6	4.5	420	11	08R1G1	08R1G1 mus musculus	603	6	4.5	475	10	09LAK7	09LAK7 arabidopsis
531	6	4.5	421	5	023852	023852 dictyostel	604	6	4.5	476	11	09O320	09O320 mus musculus
532	6	4.5	421	10	09SAD2	09SAD2 arabidopsis	605	6	4.5	476	11	09CXS6	09CXS6 mus musculus
533	6	4.5	422	5	018037	018037 caenorhabdi	606	6	4.5	476	13	09Y119	09Y119 agkistrodon
534	6	4.5	424	13	0919Y8	0919Y8 bonasa umbe	607	6	4.5	476	16	08YRZ8	08YRZ8 anabaena sp
535	6	4.5	425	2	09JPP7	09JPP7 neisseria m	608	6	4.5	479	16	091006	091006 pseudomonas
536	6	4.5	426	2	09JPH2	09JPH2 neisseria m	609	6	4.5	480	2	09JPP4	09JPP4 neisseria m
537	6	4.5	426	2	09JPH2	09JPH2 neisseria m	610	6	4.5	481	2	09JPP9	09JPP9 neisseria m
538	6	4.5	426	4	060410	060410 homo sapien	611	6	4.5	482	13	09PYK9	09PYK9 agkistrodon
539	6	4.5	426	4	015243	015243 homo sapien	612	6	4.5	482	11	09PNB0	09PNB0 mus musculus
540	6	4.5	426	4	016660	016660 homo sapien	613	6	4.5	484	11	09EY78	09EY78 mus musculus
541	6	4.5	426	16	09RUD9	09RUD9 delinococcus	614	6	4.5	485	2	09JPP5	09JPP5 neisseria m
542	6	4.5	427	2	09JPI1	09JPI1 neisseria m	615	6	4.5	486	11	08RT07	08RT07 mus musculus
543	6	4.5	427	2	09JPE1	09JPE1 neisseria m	616	6	4.5	487	17	08RT25	08RT25 methanosarc
544	6	4.5	427	5	09N3E2	09N3E2 caenorhabdi	617	6	4.5	488	2	09JPD2	09JPD2 neisseria m
545	6	4.5	427	10	09LZ03	09LZ03 arabidopsis	618	6	4.5	488	10	08YVJ4	08YVJ4 arabidopsis
546	6	4.5	428	2	09JPD0	09JPD0 neisseria m	619	6	4.5	488	10	09JQW0	09JQW0 neisseria m
547	6	4.5	428	4	09JPM5	09JPM5 homo sapien	620	6	4.5	491	17	09HNT1	09HNT1 halobacteri
548	6	4.5	428	4	008265	008265 homo sapien	621	6	4.5	492	2	09JPH1	09JPH1 neisseria m
549	6	4.5	428	5	09N1N1	09N1N1 babesia mic	622	6	4.5	493	2	030495	030495 pseudomonas
550	6	4.5	429	2	09K2R4	09K2R4 neisseria g	623	6	4.5	493	9	09JMM8	09JMM8 bacterioph
551	6	4.5	429	16	09A018	09A018 streptococc	624	6	4.5	495	10	065539	065539 arabidopsis
552	6	4.5	429	17	09V1M1	09V1M1 pyrococcus	625	6	4.5	495	16	0987M6	0987M6 rhizobium l
553	6	4.5	430	16	0824W0	0824W0 salmonella	626	6	4.5	496	2	09JPP8	09JPP8 neisseria m
554	6	4.5	431	10	022616	022616 lycopersico	627	6	4.5	497	2	09JPP6	09JPP6 neisseria m
555	6	4.5	431	16	08YX30	08YX30 anabaena sp	628	6	4.5	497	16	09JTR03	09JTR03 neisseria m
556	6	4.5	432	2	P74822	P74822 sphingomon	629	6	4.5	498	10	092R09	092R09 arabidopsis
557	6	4.5	432	10	09SXF3	09SXF3 nicotiana t	630	6	4.5	499	4	096CL5	096CL5 homo sapien
558	6	4.5	432	10	09EPK5	09EPK5 nicotiana g	631	6	4.5	499	16	092EC8	092EC8 listeria in
559	6	4.5	433	10	P93351	P93351 nicotiana t	632	6	4.5	502	16	09PMZ2	09PMZ2 campylobact
560	6	4.5	434	5	017328	017328 caenorhabdi	633	6	4.5	502	16	08X4B7	08X4B7 escherichia
561	6	4.5	434	5	08T4M7	08T4M7 giardia lam	634	6	4.5	505	16	09KPD2	09KPD2 vibrio chol
562	6	4.5	435	10	08S3N2	08S3N2 capsicum an	635	6	4.5	505	16	09KPD2	09KPD2 vibrio chol
563	6	4.5	436	16	08ED46	08ED46 yersinia pe	636	6	4.5	506	16	09CFL4	09CFL4 lactococcus
564	6	4.5	438	12	065165	065165 african swi	637	6	4.5	507	16	08XJ35	08XJ35 clostridium
565	6	4.5	439	16	067198	067198 aquifex aeo	638	6	4.5	514	15	09DLF6	09DLF6 human immun
566	6	4.5	440	5	09VW75	09VW75 drosophila	639	6	4.5	514	15	09DLF5	09DLF5 human immun
567	6	4.5	440	12	0911G5	0911G5 white spot	640	6	4.5	514	15	09DLF4	09DLF4 human immun
568	6	4.5	440	16	08XB35	08XB35 escherichia	641	6	4.5	514	15	09DLF3	09DLF3 human immun
569	6	4.5	444	3	08TGC2	08TGC2 aspergillus	642	6	4.5	514	15	09DLF2	09DLF2 human immun
570	6	4.5	444	5	026895	026895 trypanosoma	643	6	4.5	514	15	09DLF1	09DLF1 human immun
571	6	4.5	445	5	09NF64	09NF64 caenorhabdi	644	6	4.5	514	15	09DLF1	09DLF1 human immun
572	6	4.5	445	5	020022	020022 caenorhabdi	645	6	4.5	515	15	09DLF0	09DLF0 human immun
573	6	4.5	445	16	08R693	08R693 fusobacteri	646	6	4.5	515	17	08THA5	08THA5 methanosarc
574	6	4.5	446	10	09ZRO7	09ZRO7 arabidopsis	647	6	4.5	518	11	09CWP6	09CWP6 mus musculus
575	6	4.5	446	10	09FKH1	09FKH1 arabidopsis	648	6	4.5	521	5	08SYR4	08SYR4 drosophila
576	6	4.5	446	16	033076	033076 mycobacteri	649	6	4.5	522	11	09CXE1	09CXE1 mus musculus
577	6	4.5	447	16	0988E5	0988E5 rhizobium l	650	6	4.5	523	15	085477	085477 rous sarcom
578	6	4.5	449	16	097F97	097F97 clostridium	651	6	4.5	525	2	09KHC7	09KHC7 streptomyce
579	6	4.5	449	16	09ZS43	09ZS43 rhizobium m	652	6	4.5	525	16	08RHY0	08RHY0 fusobacteri
580	6	4.5	450	4	09BRW2	09BRW2 homo sapien	653	6	4.5	526	8	09T9G6	09T9G6 pupa strigo
581	6	4.5	453	10	P92981	P92981 arabidopsis	654	6	4.5	526	15	064993	064993 rous sarcom
582	6	4.5	453	10	08YV23	08YV23 arabidopsis	655	6	4.5	526	15	064994	064994 rous sarcom
583	6	4.5	454	10	022554	022554 arabidopsis	656	6	4.5	526	15	093080	093080 rous sarcom
584	6	4.5	454	10	0946G4	0946G4 chlamydomon	657	6	4.5	526	15	060567	060567 rous sarcom
585	6	4.5	454	10	0944U2	0944U2 arabidopsis	658	6	4.5	526	15	092806	092806 rous sarcom
586	6	4.5	454	10	004583	004583 arabidopsis	659	6	4.5	526	15	007461	007461 rous sarcom
587	6	4.5	454	16	08T6Q1	08T6Q1 listeria mo	660	6	4.5	527	16	09ZAH8	09ZAH8 staphylococ
588	6	4.5	455	10	004215	004215 arabidopsis	661	6	4.5	528	6	08WM97	08WM97 macaca fasc
589	6	4.5	457	16	092G54	092G54 rickettsia	662	6	4.5	529	12	09YZA2	09YZA2 simlan para
590	6	4.5	459	10	09PE88	09PE88 xyella fas	663	6	4.5	530	5	0917L5	0917L5 drosophila
591	6	4.5	459	10	09SGY0	09SGY0 arabidopsis	664	6	4.5	531	10	082416	082416 papaver som
592	6	4.5	459	12	09MT19	09MT19 human herpe	665	6	4.5	532	9	09K3U5	09K3U5 streptomyce
593	6	4.5	459	12	09OJ29	09OJ29 human herpe	666	6	4.5	532	5	044117	044117 drosophila
594	6	4.5	461	5	09N8A1	09N8A1 plasmodium	667	6	4.5	534	10	081790	081790 arabidopsis
595	6	4.5	461	16	08UGG4	08UGG4 agrobacteri	668	6	4.5	534	12	083516	083516 mutayama vi
596	6	4.5	462	2	091U09	091U09 escherichia	669	6	4.5	535	4	096EB1	096EB1 homo sapien
597	6	4.5	463	10	09LJ51	09LJ51 arabidopsis	670	6	4.5	535	15	092957	092957 rous sarcom
598	6	4.5	464	5	09TYE5	09TYE5 trypanosoma	671	6	4.5	536	5	09TZL9	09TZL9 caenorhabdi
599	6	4.5	469	16	09ZM28	09ZM28 rhizobium m	672	6	4.5				
600	6	4.5	473	10	09LSG2	09LSG2 arabidopsis	673	6	4.5				

674	6	4.5	536	12	090PR2	090pr2 newcastle d	747	6	4.5	551	12	080T25	080t25 newcastle d
675	6	4.5	536	12	09E7M0	09e7m0 chimeric hu	748	6	4.5	551	12	080T24	080t24 newcastle d
676	6	4.5	536	12	09E7M8	09e7m8 chimeric hu	749	6	4.5	551	12	080T23	080t23 newcastle d
677	6	4.5	536	12	084845	084845 avian param	750	6	4.5	551	12	080T22	080t22 newcastle d
678	6	4.5	537	6	09WZU5	09wzu5 sus scrofa	751	6	4.5	551	12	080T21	080t21 newcastle d
679	6	4.5	537	10	09SL09	09sl09 aradidopsis	752	6	4.5	551	12	080T20	080t20 newcastle d
680	6	4.5	538	12	083614	083614 mumps virus	753	6	4.5	551	12	080T19	080t19 newcastle d
681	6	4.5	538	12	083653	083653 mumps virus	754	6	4.5	551	12	080T18	080t18 newcastle d
682	6	4.5	538	12	09WAE9	09wae9 mumps virus	755	6	4.5	551	12	080T17	080t17 newcastle d
683	6	4.5	538	12	09WAF1	09waf1 mumps virus	756	6	4.5	551	12	080T16	080t16 newcastle d
684	6	4.5	538	12	09WAF4	09waf4 mumps virus	757	6	4.5	551	12	080T15	080t15 newcastle d
685	6	4.5	538	12	09WAF6	09waf6 mumps virus	758	6	4.5	551	12	080T14	080t14 newcastle d
686	6	4.5	538	12	09WAF9	09waf9 mumps virus	759	6	4.5	551	12	080T13	080t13 newcastle d
687	6	4.5	538	12	09WAG2	09wag2 mumps virus	760	6	4.5	551	12	080T12	080t12 newcastle d
688	6	4.5	538	12	09WAG5	09wag5 mumps virus	761	6	4.5	551	16	08YPR8	08ypr8 arabidopsis
689	6	4.5	538	12	09WAG7	09wag7 mumps virus	762	6	4.5	553	12	056771	056771 neurospora
690	6	4.5	538	12	09WAG9	09wag9 mumps virus	763	6	4.5	553	12	083849	083849 newcastle d
691	6	4.5	538	12	09WAG1	09wag1 mumps virus	764	6	4.5	553	12	083847	083847 newcastle d
692	6	4.5	538	12	09WAG3	09wag3 mumps virus	765	6	4.5	553	12	090339	090339 newcastle d
693	6	4.5	538	12	09WAG6	09wag6 mumps virus	766	6	4.5	553	12	090339	090339 newcastle d
694	6	4.5	538	12	09WAG8	09wag8 mumps virus	767	6	4.5	553	12	09WD11	09wd11 newcastle d
695	6	4.5	538	12	09WAG10	09wag10 mumps virus	768	6	4.5	553	12	09JIC8	09jic8 newcastle d
696	6	4.5	538	12	09WAG12	09wag12 mumps virus	769	6	4.5	553	12	09JIC8	09jic8 newcastle d
697	6	4.5	538	12	09JAE8	09jae8 mumps virus	770	6	4.5	553	12	09DLD4	09dld4 newcastle d
698	6	4.5	538	12	09JAE9	09jae9 mumps virus	771	6	4.5	553	12	09WMT7	09wmt7 newcastle d
699	6	4.5	538	12	09JAE7	09jae7 mumps virus	772	6	4.5	553	12	083854	083854 newcastle d
700	6	4.5	538	12	09JAE6	09jae6 mumps virus	773	6	4.5	553	12	083855	083855 newcastle d
701	6	4.5	538	12	09JAE5	09jae5 mumps virus	774	6	4.5	553	12	083856	083856 newcastle d
702	6	4.5	538	12	09JAE4	09jae4 mumps virus	775	6	4.5	553	12	090362	090362 newcastle d
703	6	4.5	538	12	09JAE3	09jae3 mumps virus	776	6	4.5	553	12	09WLE2	09wle2 newcastle d
704	6	4.5	538	12	09JAE2	09jae2 mumps virus	777	6	4.5	553	12	09WLE1	09wle1 newcastle d
705	6	4.5	538	12	09JAE1	09jae1 mumps virus	778	6	4.5	553	12	09WLE1	09wle1 newcastle d
706	6	4.5	538	12	09JAE0	09jae0 mumps virus	779	6	4.5	553	12	09WLE1	09wle1 newcastle d
707	6	4.5	538	12	09JAD8	09jad8 mumps virus	780	6	4.5	553	12	09WLE1	09wle1 newcastle d
708	6	4.5	538	12	09JAD9	09jad9 mumps virus	781	6	4.5	553	12	09WLE1	09wle1 newcastle d
709	6	4.5	538	12	09JAD8	09jad8 mumps virus	782	6	4.5	553	12	09WLE1	09wle1 newcastle d
710	6	4.5	538	12	09JAD2	09jad2 mumps virus	783	6	4.5	553	12	09WLE1	09wle1 newcastle d
711	6	4.5	538	12	09JAD2	09jad2 mumps virus	784	6	4.5	553	12	09WLE1	09wle1 newcastle d
712	6	4.5	538	12	09YKB4	09ykb4 mumps virus	785	6	4.5	553	12	09WLE1	09wle1 newcastle d
713	6	4.5	538	12	09YKB3	09ykb3 mumps virus	786	6	4.5	553	12	09WLE1	09wle1 newcastle d
714	6	4.5	538	12	09YJW8	09yjw8 mumps virus	787	6	4.5	553	12	09WLE1	09wle1 newcastle d
715	6	4.5	538	12	09YJW8	09yjw8 mumps virus	788	6	4.5	553	12	09WLE1	09wle1 newcastle d
716	6	4.5	538	12	09YJW8	09yjw8 mumps virus	789	6	4.5	553	12	09WLE1	09wle1 newcastle d
717	6	4.5	538	12	09YJW8	09yjw8 mumps virus	790	6	4.5	553	12	09WLE1	09wle1 newcastle d
718	6	4.5	538	12	09YJW8	09yjw8 mumps virus	791	6	4.5	553	12	09WLE1	09wle1 newcastle d
719	6	4.5	538	12	09YJW8	09yjw8 mumps virus	792	6	4.5	553	12	09WLE1	09wle1 newcastle d
720	6	4.5	538	12	09YJW8	09yjw8 mumps virus	793	6	4.5	553	12	09WLE1	09wle1 newcastle d
721	6	4.5	538	12	09YJW8	09yjw8 mumps virus	794	6	4.5	553	12	09WLE1	09wle1 newcastle d
722	6	4.5	540	6	09SK25	09sk25 macaca fasc	795	6	4.5	553	12	09WLE1	09wle1 newcastle d
723	6	4.5	540	8	09MT29	09mt29 solanum tub	796	6	4.5	553	12	09WLE1	09wle1 newcastle d
724	6	4.5	541	12	P89039	P89039 paramyxovir	797	6	4.5	553	12	09WLE1	09wle1 newcastle d
725	6	4.5	542	11	09JJI0	09jji0 raltus norv	798	6	4.5	553	12	09WLE1	09wle1 newcastle d
726	6	4.5	545	15	086362	086362 rous sarcom	799	6	4.5	553	12	09WLE1	09wle1 newcastle d
727	6	4.5	546	2	093KDB	093kdb chlorogloeo	800	6	4.5	553	12	09WLE1	09wle1 newcastle d
728	6	4.5	546	2	086363	086363 rous sarcom	801	6	4.5	553	12	09WLE1	09wle1 newcastle d
729	6	4.5	547	2	086363	086363 rous sarcom	802	6	4.5	553	12	09WLE1	09wle1 newcastle d
730	6	4.5	547	2	086363	086363 rous sarcom	803	6	4.5	553	12	09WLE1	09wle1 newcastle d
731	6	4.5	547	16	09A2R2	09a2r2 cauliobacter	804	6	4.5	553	12	09WLE1	09wle1 newcastle d
732	6	4.5	548	3	000357	000357 cociliobol	805	6	4.5	553	12	09WLE1	09wle1 newcastle d
733	6	4.5	548	10	09LUF7	09luf7 aradidopsis	806	6	4.5	553	12	09WLE1	09wle1 newcastle d
734	6	4.5	549	16	092DM5	092dm5 rhizobium m	807	6	4.5	553	12	09WLE1	09wle1 newcastle d
735	6	4.5	551	3	098975	098975 saccharomyc	808	6	4.5	553	12	09WLE1	09wle1 newcastle d
736	6	4.5	551	12	090MH1	090mh1 simian para	809	6	4.5	553	12	09WLE1	09wle1 newcastle d
737	6	4.5	551	12	09E7N2	09e7n2 human para	810	6	4.5	553	12	09WLE1	09wle1 newcastle d
738	6	4.5	551	12	09DUM4	09dum4 porcine rub	811	6	4.5	553	12	09WLE1	09wle1 newcastle d
739	6	4.5	551	12	08VIO7	08vio7 newcastle d	812	6	4.5	553	12	09WLE1	09wle1 newcastle d
740	6	4.5	551	12	08VIO6	08vio6 newcastle d	813	6	4.5	553	12	09WLE1	09wle1 newcastle d
741	6	4.5	551	12	08VIO5	08vio5 newcastle d	814	6	4.5	553	12	09WLE1	09wle1 newcastle d
742	6	4.5	551	12	08VIO4	08vio4 newcastle d	815	6	4.5	553	12	09WLE1	09wle1 newcastle d
743	6	4.5	551	12	08VIO3	08vio3 newcastle d	816	6	4.5	553	12	09WLE1	09wle1 newcastle d
744	6	4.5	551	12	08VIO2	08vio2 newcastle d	817	6	4.5	553	12	09WLE1	09wle1 newcastle d
745	6	4.5	551	12	08VIO0	08vio0 newcastle d	818	6	4.5	553	12	09WLE1	09wle1 newcastle d
746	6	4.5	551	12	08VIO9	08vio9 newcastle d	819	6	4.5	553	12	09WLE1	09wle1 newcastle d

820	6	4.5	604	3	Q9US18	Q9us18 schizosacch	893	6	4.5	695	5	Q9UE83	Q9ue83 drosophila
821	6	4.5	605	16	Q9X1N4	Q9x1n4 thermotoga	894	6	4.5	695	5	Q9UE82	Q9ue82 drosophila
822	6	4.5	606	15	Q03803	Q03803 avian leuco	895	6	4.5	695	5	Q9UE81	Q9ue81 drosophila
823	6	4.5	613	5	Q22970	Q22970 caenorhabd	896	6	4.5	695	5	Q9UE80	Q9ue80 hirtodrosop
824	6	4.5	613	12	Q80TE5	Q80te5 white spot	897	6	4.5	695	5	Q9UE89	Q9ue89 drosophila
825	6	4.5	614	4	Q9Y2D8	Q9y2d8 homo sapien	898	6	4.5	695	5	Q9UE87	Q9ue87 chymomyza a
826	6	4.5	614	4	Q9Y2D8	Q9y2d8 homo sapien	899	6	4.5	696	2	Q54201	Q54201 streptomyc
827	6	4.5	619	2	Q9F5M0	Q9f5m0 zymomonas m	900	6	4.5	696	5	Q8T057	Q8t057 drosophila
828	6	4.5	635	10	Q8SOV5	Q8sov5 oryza sativ	901	6	4.5	696	16	Q98K28	Q98k28 rhizobium 1
829	6	4.5	637	5	Q9BPN8	Q9bpn8 caenorhabd	902	6	4.5	698	13	Q9W724	Q9w724 cyprinus ca
830	6	4.5	638	3	Q07442	Q07442 saccharomyc	903	6	4.5	703	3	Q9L3P4	Q9l3p4 uncultured
831	6	4.5	644	4	Q14985	Q14985 homo sapien	904	6	4.5	703	2	Q9FAP9	Q9fap9 uncultured
832	6	4.5	646	16	Q8X086	Q8x086 ralstonia s	905	6	4.5	703	2	Q9EXP9	Q9exp9 bacteroides
833	6	4.5	647	10	Q9AVV1	Q9avv1 pharbitlis n	906	6	4.5	705	10	Q9F7F3	Q9f7f3 arbidops
834	6	4.5	649	5	Q9GTF7	Q9gtf7 drosophila	907	6	4.5	706	10	Q9FV97	Q9fv97 arbidops
835	6	4.5	650	4	Q9H8I6	Q9h8i6 homo sapien	908	6	4.5	707	16	Q9EWD1	Q9ewd1 escherichia
836	6	4.5	652	10	Q9FH34	Q9fh34 arbidops	909	6	4.5	713	10	Q9LWT2	Q9lwt2 arbidops
837	6	4.5	652	10	Q9FG33	Q9fg33 arbidops	910	6	4.5	715	10	Q9SH86	Q9sh86 arbidops
838	6	4.5	653	3	Q12171	Q12171 saccharomyc	911	6	4.5	720	5	Q960B4	Q960b4 drosophila
839	6	4.5	658	11	Q9D4H3	Q9d4h3 mus musculu	912	6	4.5	721	16	Q92EJ0	Q92ej0 listeria in
840	6	4.5	659	15	Q9DLF9	Q9dlf9 human immun	913	6	4.5	722	12	Q9LTV9	Q9ltv9 lupala herp
841	6	4.5	659	15	Q9DLF8	Q9dlf8 human immun	914	6	4.5	726	3	Q9UW98	Q9uw98 trichophyto
842	6	4.5	659	15	Q9DLF7	Q9dlf7 human immun	915	6	4.5	726	10	Q9LWK3	Q9lwk3 arbidops
843	6	4.5	662	11	Q91WV6	Q91wv6 mus musculu	916	6	4.5	726	17	Q8TWD8	Q8twd8 methanosarc
844	6	4.5	663	5	Q62446	Q62446 caenorhabd	917	6	4.5	730	8	Q98698	Q98698 exacum affi
845	6	4.5	663	9	Q94MS4	Q94ms4 bacterioph	918	6	4.5	734	13	Q9W6K2	Q9w6k2 xenopus lae
846	6	4.5	664	4	Q92541	Q92541 homo sapien	919	6	4.5	735	5	Q8SKC4	Q8skc4 drosophila
847	6	4.5	664	10	Q8VZP8	Q8vzp8 arbidops	920	6	4.5	744	10	Q9MB58	Q9mb58 arbidops
848	6	4.5	670	2	Q8VNN7	Q8vnn7 escherichia	921	6	4.5	744	10	Q9SP17	Q9sp17 arbidops
849	6	4.5	670	16	Q8RAU5	Q8rau5 thermosae	922	6	4.5	745	10	Q9LLI3	Q9lli3 arbidops
850	6	4.5	674	4	Q96KR2	Q96kr2 homo sapien	923	6	4.5	745	10	Q9LZES	Q9lzes arbidops
851	6	4.5	678	2	Q9JN46	Q9jnn46 rhodobacter	924	6	4.5	746	10	Q9LTV8	Q9lvt8 arbidops
852	6	4.5	678	16	Q9HX82	Q9hxr2 pseudomonas	925	6	4.5	748	5	Q9VTE7	Q9vte7 drosophila
853	6	4.5	679	9	Q9MBK4	Q9mbk4 bacterioph	926	6	4.5	752	10	Q9LTB0	Q9ltb0 arbidops
854	6	4.5	680	16	Q97DC7	Q97dc7 clostridium	927	6	4.5	752	12	Q9L3V5	Q9l3v5 baboon gamm
855	6	4.5	681	4	Q92624	Q92624 homo sapien	928	6	4.5	755	2	Q9LTV9	Q9lvt9 arbidops
856	6	4.5	681	15	Q9YWM1	Q9ywm1 gibbon ape	929	6	4.5	755	2	Q93M42	Q93m42 streptococ
857	6	4.5	683	16	Q8XU96	Q8xu96 ralstonia s	930	6	4.5	762	2	Q87714	Q87714 caulobacter
858	6	4.5	685	6	Q9TTS5	Q9tte5 bos taurus	931	6	4.5	762	2	Q9RM63	Q9rm63 myroides od
859	6	4.5	685	15	Q9YWM3	Q9ywm3 gibbon ape	932	6	4.5	762	16	Q9AB02	Q9ab02 caulobacter
860	6	4.5	685	15	Q9YWM2	Q9ywm2 gibbon ape	933	6	4.5	763	4	Q9UPP4	Q9upp4 homo sapien
861	6	4.5	686	5	Q9GTG1	Q9gtg1 caenorhabd	934	6	4.5	765	3	Q94672	Q94672 schizosacch
862	6	4.5	690	5	Q9VFC5	Q9vfc5 drosophila	935	6	4.5	771	11	Q63693	Q63693 rattus norv
863	6	4.5	692	5	Q9M1H4	Q9m1h4 drosophila	936	6	4.5	772	4	Q9Y468	Q9y468 homo sapien
864	6	4.5	695	5	Q9NJB3	Q9njb3 drosophila	937	6	4.5	772	11	Q62711	Q62711 rattus norv
865	6	4.5	695	5	Q9NJB2	Q9njb2 drosophila	938	6	4.5	773	5	Q9V475	Q9v475 drosophila
866	6	4.5	695	5	Q9NJB1	Q9njb1 drosophila	939	6	4.5	774	16	Q9K2S2	Q9k2s2 bacillus su
867	6	4.5	695	5	Q9NCB3	Q9ncb3 drosophila	940	6	4.5	776	10	Q9FTJ3	Q9ftj3 arbidops
868	6	4.5	695	5	Q9NCB1	Q9ncb1 drosophila	941	6	4.5	776	10	Q9FTJ3	Q9ftj3 arbidops
869	6	4.5	695	5	Q9NCB0	Q9ncb0 drosophila	942	6	4.5	779	5	Q9W0U8	Q9w0u8 arbidops
870	6	4.5	695	5	Q9NCB6	Q9ncb6 drosophila	943	6	4.5	787	10	Q9C507	Q9c507 arbidops
871	6	4.5	695	5	Q9NCB5	Q9ncb5 drosophila	944	6	4.5	789	5	Q61082	Q61082 trypanosoma
872	6	4.5	695	5	Q9NCB4	Q9ncb4 drosophila	945	6	4.5	790	16	Q8ZNI3	Q8zni3 salmonella
873	6	4.5	695	5	Q9NCB1	Q9ncb1 drosophila	946	6	4.5	795	10	Q9M366	Q9m366 arbidops
874	6	4.5	695	5	Q9N602	Q9n602 drosophila	947	6	4.5	795	16	Q97EA9	Q97ea9 clostridium
875	6	4.5	695	5	Q9N2R9	Q9n2r9 drosophila	948	6	4.5	804	4	Q9UPQ3	Q9upq3 homo sapien
876	6	4.5	695	5	Q9N2R8	Q9n2r8 drosophila	949	6	4.5	807	5	Q21895	Q21895 caenorhabd
877	6	4.5	695	5	Q9N2R6	Q9n2r6 drosophila	950	6	4.5	807	16	Q985D7	Q985d7 rhizobium 1
878	6	4.5	695	5	Q9N2R5	Q9n2r5 drosophila	951	6	4.5	815	3	Q59744	Q59744 schizosacch
879	6	4.5	695	5	Q9N2R4	Q9n2r4 drosophila	952	6	4.5	815	3	Q961J4	Q961j4 drosophila
880	6	4.5	695	5	Q9N2R6	Q9n2r6 drosophila	953	6	4.5	823	10	Q9ZVT8	Q9zvt8 arbidops
881	6	4.5	695	5	Q95WB5	Q95wb5 drosophila	954	6	4.5	828	2	Q51843	Q51843 cellulomona
882	6	4.5	695	5	Q95WB3	Q95wb3 drosophila	955	6	4.5	833	11	Q921O6	Q921o6 mus musculu
883	6	4.5	695	5	Q9NCA2	Q9ncb2 drosophila	956	6	4.5	833	11	Q9QUM7	Q9qum7 mus musculu
884	6	4.5	695	5	Q9NCA3	Q9ncb3 drosophila	957	6	4.5	835	12	Q71304	Q71304 woolly monk
885	6	4.5	695	5	Q9NCA4	Q9ncb4 drosophila	958	6	4.5	840	4	Q9H4M1	Q9h4m1 homo sapien
886	6	4.5	695	5	Q9NCA7	Q9ncb7 drosophila	959	6	4.5	840	4	Q8TCV0	Q8tcv0 homo sapien
887	6	4.5	695	5	Q9NCA8	Q9ncb8 drosophila	960	6	4.5	847	10	Q9STF7	Q9stf7 arbidops
888	6	4.5	695	5	Q96441	Q96441 drosophila	961	6	4.5	849	10	Q9SHE4	Q9she4 arbidops
889	6	4.5	695	5	Q9UE88	Q9ue88 drosophila	962	6	4.5	853	5	Q9VDS5	Q9vds5 arbidops
890	6	4.5	695	5	Q9UE87	Q9ue87 drosophila	963	6	4.5	853	10	Q9FNU6	Q9fnu6 arbidops
891	6	4.5	695	5	Q9UE86	Q9ue86 drosophila	964	6	4.5	853	12	Q93123	Q93123 human calic
892	6	4.5	695	5	Q9UE85	Q9ue85 drosophila	965	6	4.5	859	16	P94181	P94181 anabaena sp

966	6	4.5	877	5	023853	023853 dictyostell
967	6	4.5	880	16	08xt42	08xt42 clostridium
968	6	4.5	882	11	09d44	09d44 mus musculus
969	6	4.5	884	5	061543	061543 drosophila
970	6	4.5	884	5	09vev9	09vev9 drosophila
971	6	4.5	886	16	09cbv7	09cbv7 mycobacteri
972	6	4.5	887	10	09fn93	09fn93 arabidopsis
973	6	4.5	892	12	09jfp0	09jfp0 vaccinia vi
974	6	4.5	895	5	095208	095208 trypanosoma
975	6	4.5	901	16	08rcb4	08rcb4 drosophila
976	6	4.5	903	5	095t22	095t22 thermococ
977	6	4.5	903	5	09vc87	09vc87 drosophila
978	6	4.5	904	6	09nxz1	09nxz1 homo sapien
979	6	4.5	906	6	08sof7	08sof7 rhinopoma h
980	6	4.5	914	5	09n4s5	09n4s5 caenorhabdi
981	6	4.5	914	5	08w09	08w09 drosophila
982	6	4.5	916	16	08u442	08u442 agrobacteri
983	6	4.5	919	4	09hap2	09hap2 homo sapien
984	6	4.5	923	16	0915w4	0915w4 pseudomonas
985	6	4.5	928	16	092aa9	092aa9 listeria in
986	6	4.5	930	5	096165	096165 plasmodium
987	6	4.5	932	10	09lma3	09lma3 arabidopsis
988	6	4.5	940	5	076340	076340 manduca sex
989	6	4.5	942	10	09s0t8	09s0t8 arabidopsis
990	6	4.5	943	12	086545	086545 simian yiru
991	6	4.5	947	2	093fv1	093fv1 bacteroides
992	6	4.5	948	10	081211	081211 arabidopsis
993	6	4.5	953	10	08vvc7	08vvc7 arabidopsis
994	6	4.5	966	4	09c0c4	09c0c4 homo sapien
995	6	4.5	966	10	08s0z3	08s0z3 oryza sativ
996	6	4.5	967	3	008294	008294 saccharomyc
997	6	4.5	969	5	09nd19	09nd19 plasmodium
998	6	4.5	971	11	035382	035382 mus musculu
999	6	4.5	974	4	096a65	096a65 homo sapien
1000	6	4.5				

ALIGNMENTS

RESULT 1
Q8YHV9 PRELIMINARY: PRT: 100 AA.
AC Q8YHV9: 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DE Hypothetical protein BME10687.
GN BME10687.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OX NCBI_TaxID=29459;
[1]
SEQUENCE FROM N.A.
RP STRAIN-16M / ATCC 23456 / BIOTYPE 1;
RC MEDLINE-20020109; PubMed-11756688;
RA DelVecchio V.G., Kapralov V., Redkar R.J., Patra G., Mujar C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonki L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldstein E.,
RA Selkov E., Elizer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL: AE009510; AAL51868.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 100 AA; 11110 MW; 787FER2162163843 CRC64;

Query Match 6.0%; Score 8; DB 16; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AFLASCA 21
Db 71 AFLASCA 78

RESULT 2
Q92RM9 PRELIMINARY: PRT: 120 AA.
AC Q92RM9: 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE Putative FOPI ATP synthase, subunit I transmembrane protein.
GN ATP OR R00834 OR SMC00872.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OX NCBI_TaxID=382;
[1]
SEQUENCE FROM N.A.
RP STRAIN-1021;
RC MEDLINE-21396507; PubMed-11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetle D., Puehler A., Purrelle B., Rampeger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,
RT "Analysis of the chromosome sequence of the legume symbiont
RL Sinorhizobium meliloti strain 1021.";
DR Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
KW EMBL: AL591785; CAC45406.1; -
SQ SEQUENCE 120 AA; 12589 MW; 87210F95FA905978 CRC64;

Query Match 6.0%; Score 8; DB 16; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 AAMKLSE 68
Db 45 AAMKLSE 52

RESULT 3
Q14882 PRELIMINARY: PRT: 330 AA.
AC Q14882: 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE Mucin (Fragment).
GN MUC5B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RP TISSUE=TRACHEOBRONCHIAL MUCOSA;
RC Aubert J.;
DT Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL: X74956; CAA52911.1; -
FT NON_TER 1 330
FT NON_TER 1 330
SQ SEQUENCE 330 AA; 35555 MW; 97892E19FFC4FA50 CRC64;

Query Match 6.0%; Score 8; DB 4; Length 330;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 91 LTELTTAA 98
Db 26 LTELTTAA 33

RESULT 4
ID 014879 PRELIMINARY; PRT; 543 AA.
AC 014879;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE Mucin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TRACHEOBRONCHIAL MUCOSA;
RX MEDLINE=93343858; PubMed=7916618;
RA Dufosse J., Porchet N., Audie J.P., Guyonnet Duperrat V., Laine A.,
RA Van-Seuningen I., Marikachi S., Degand P., Aubert J.P.;
RT "Degenerate 87-base-pair tandem repeats create hydrophilic/hydrophobic
RT alternating domains in human mucin peptides mapped to 11p15";
RL Biochem. J. 293:328-337(1993).
DR EMBL; X/4370; CAA52408.1; -.
FT NON_TER 1
FT 543 543
SQ SEQUENCE 543 AA; 55197 MW; 6767A5E3518B287B CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 4; Length 543;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 91 LTELTTAA 98
|||||
DB 174 LTELTTAA 181

RESULT 5
ID 014881 PRELIMINARY; PRT; 622 AA.
AC 014881;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Mucin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TRACHEOBRONCHIAL MUCOSA;
RX Desseyn J.L., Guyonnet-Duperrat V., Porchet N., Aubert J.P., Laine A.;
RT "Human mucin gene MUC5B: the 10.7 kb large central exon encodes
RT various alternate subdomains resulting in a super repeat";
RT Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X/4955; CAA52910.1; -.
FT NON_TER 1
FT 622 622
SQ SEQUENCE 622 AA; 61786 MW; 4FC85A52F50D57E5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 4; Length 622;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 91 LTELTTAA 98
|||||
DB 223 LTELTTAA 230

RESULT 6

Q9NVE4
ID 09NVE4 PRELIMINARY; PRT; 716 AA.
AC 09NVE4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE Cervical mucin MUC5B (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ho S.B., Toribara N.W., Anway R.E., Spurr-Michaud S.J., Shekels L.L.,
RA Keutmann H.T., Hill J.A., Gipson I.K.;
RT "Expression cloning of human cervical proteins using an antibody to
RT cervical mucus";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF253321; AAF64523.1; -.
FT NON_TER 1
FT 716 716
SQ SEQUENCE 716 AA; 71194 MW; 6F3F20E7512289F3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 4; Length 716;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 91 LTELTTAA 98
|||||
DB 483 LTELTTAA 490

RESULT 7
ID Q8UZD5 PRELIMINARY; PRT; 864 AA.
AC Q8UZD5;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE BALF4
OS cercopithicline herpesvirus 15.
OC viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=104228;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX Franken M., Devereux O., Rosenzweig M., Anis B., Kieff E., Wang F.;
RT "Comparative analysis identifies conserved tumor necrosis factor
RT receptor-associated factor 3 binding sites in the human and simian
RT Epstein-Barr virus oncogene LMP1";
RL J. Virol. 70:7819-7826(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=99412410; PubMed=10482645;
RX Rivalier P., Quink C., Wang F.;
RT "Strong selective pressure for evolution of an Epstein-Barr virus
RT LMP2B homologue in the rhesus lymphocryptovirus";
RL J. Virol. 73:8867-8872(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=20304984; PubMed=10846073;
RX Jiang H., Cho Y.G., Wang F.;
RT "Structural, functional, and genetic comparisons of Epstein-Barr virus
RT lymphocryptovirus";
RT J. Virol. 74:5921-5932(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;

RX MEDLINE-20440633; PubMed-10970361;
RAO P., JIANG H., WANG F.;
RT Cloning of the Rhesus lymphocryptovirus viral capsid antigen and
RT Epstein-Barr virus-encoded small RNA homologues and use in diagnosis
RT J. Clin. Microbiol. 38:3219-3225(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN-LCL8664;
RX MEDLINE-21602573; PubMed-11739708;
RA Rivallier P., JIANG H., CHO Y.-G., QUINK C., WANG F.;
RT "Complete Nucleotide Sequence of the Rhesus lymphocryptovirus: Genetic
RT Validation for an Epstein-Barr Virus Animal Model.";
RL J. Virol. 76:421-426(2002).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN-LCL8664;
RX Moghaddam A., KOCH J., ANNIS B., WANG F.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN-LCL8664;
RX Moghaddam A., ANNIS B., WANG F.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN-LCL8664;
RX Rivallier P., QUINK C., WANG F.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [9]
RP SEQUENCE FROM N.A.
RC STRAIN-LCL8664;
RA JIANG H., WANG F.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [10]
RP SEQUENCE FROM N.A.
RC STRAIN-LCL8664;
RX Rao P.V., JIANG H., WANG F.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [11]
RP SEQUENCE FROM N.A.
RC STRAIN-LCL8664;
RA Rivallier P., JIANG H., CHO Y.-G., QUINK C., WANG F.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL:AT037858; AAK95476.1; -;
DR InterPro: IPR000234; Glycoprot. B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR PRODOM: PD000693; Glycoprot_B; 1.
SQ SEQUENCE 864 AA; 96646 MW; 89637230C484F5BD CRC64;
Query Match 6.0%; Score 8; DB 12; Length 864;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 91 LTELTTAA 98
DB 396 LTELTTAA 403
RESULT 8
ID 099552 PRELIMINARY; PRT; 3570 AA.
AC 099552;
DT 01-MAY-1997 (TREMblrel. 03, Created)
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE MUC5B (Fragment).
GN MUC5B
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
[1]

RP SEQUENCE FROM N.A.
RC TISSUE-PLACENTA;
RX MEDLINE-9716151; PubMed-9013550;
RA Desseyn J.L., Guyonnet-Duperrat V., Porchet N., Aubert J.P., Laine A.;
RT "Human mucin gene MUC5B, the 10.7 kb large central exon encodes
RT various alternate subdomains resulting in a super-repeat. Structural
RT evidence for a 11p15.5 gene family.";
RL J. Biol. Chem. 272:3168-3176(1997).
DR EMBL: 272496; CAA96577.1; -;
FT NON_TER 1 3570
FT TER 1 3570
SQ SEQUENCE 3570 AA; 361019 MW; DE04E4D72579312 CRC64;
Query Match 6.0%; Score 8; DB 4; Length 3570;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 91 LTELTTAA 98
DB 3008 LTELTTAA 3015
RESULT 9
ID 097F31 PRELIMINARY; PRT; 64 AA.
AC 097F31;
DT 01-OCT-2001 (TREMblrel. 18, Created)
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Uncharacterized protein, possibly involved in thiamine
DE biosynthesis.
DE CAC2924.
GN Clostridium acetobutylicum.
OS Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_Taxid=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE-21359325; PubMed-11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hiltl J., Wolf Y.I.,
RA Tatunov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL: AE007789; AAK80866.1; -;
DR InterPro: IPR003749; THIS.
DR Pfam: PF02597; DUF170; 1.
DR Complete proteome.
SQ SEQUENCE 64 AA; 7296 MW; 2F470F2896606E65 CRC64;
Query Match 5.2%; Score 7; DB 16; Length 64;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 64 KLSSESK 70
DB 48 KLSSESK 54
RESULT 10
ID 08YFHS PRELIMINARY; PRT; 132 AA.
AC 08YFHS;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE ATP synthase.
GN BMEI1547.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

```

OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-16M / ATCC 23456 / BIOTYPE 1;
RX MEDLINE=20020109; PubMed=11756688;
RA Delvecchio V.G., Kapatal V., Redkar R.J., Patra G., Mijer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrpides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL: AE009590; AAL52728.1; -.
KW Complete proteome.
SQ SEQUENCE 132 AA; 14090 MW; 011B830E0911528 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 AMKLSE 68
Db 62 AMKLSE 68
|||||

RESULT 11
08YIT4 PRELIMINARY; PRT: 174 AA.
ID 08YIT4;
AC 08YIT4;
DT 01-MAR-2002 (TRENBLREL. 20, Created)
DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Deoxyuridine 5'-triphosphate nucleotidohydrolase (EC 3.6.1.23).
GN BMEI0358.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-16M / ATCC 23456 / BIOTYPE 1;
RX MEDLINE=20020109; PubMed=11756688;
RA Delvecchio V.G., Kapatal V., Redkar R.J., Patra G., Mijer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrpides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL: AE009478; AAL51539.1; -.
DR InterPro: IPR001428; DeoxyUTPase.
DR Pfam: PF00692; dUTPase; 1.
DR ProDom: PD000946; DeoxyUTPase; 1.
DR TIGRFAMs: TIGR00576; dUT; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 174 AA; 18582 MW; 66588CEACC457E9C CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 AKISETA 77
Db 156 AKISETA 162
|||||

RESULT 12
09KKLO PRELIMINARY; PRT: 175 AA.
ID 09KKLO;
AC 09KKLO;

```

```

DT 01-OCT-2000 (TRENBLREL. 15, Created)
DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Purine-binding chemotaxis protein Chem.
GN VCA1094.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Yamatchevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen vibrio
RT cholerae";
RL Nature 406:477-483(2000).
DR EMBL: AE004434; AAP69686.1; -.
DR TIGR: VCA1094; -.
DR InterPro: IPR002545; Chem.
DR Pfam: PF01584; Chem; 1.
DR SMART: SM00260; Chem; 1.
KW Complete proteome.
SQ SEQUENCE 175 AA; 19068 MW; 8BC1AE790E2ED333 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 SANLPTS 38
Db 3 SANLPTS 9
|||||

RESULT 13
0975X8 PRELIMINARY; PRT: 185 AA.
ID 0975X8;
AC 0975X8;
DT 01-DEC-2001 (TRENBLREL. 19, Created)
DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Hypothetical protein ST0300.
GN ST0300.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=11955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Anai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kishida M., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain 7.";
RL DNA Res. 8:123-140(2001).
DR EMBL: AP000982; BAB55270.1; -.
DR InterPro: IPR002783; Adenylate_cyc.
DR Pfam: PF01928; Adenylate_cyc_2; 1.
DR ProDom: PD009560; Adenylate_cyc; 1.
DR TIGRFAMs: TIGR00318; cyab; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 185 AA; 21980 MW; CCCF829805FC0A09 CRC64;

```

Query Match
Best Local Similarity 100.0%; Score 7; DB 17; Length 185;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 KLSSEK 70
Db 72 KLSSEK 78

RESULT 14
O8YGD7
ID 08YGD7 PRELIMINARY; PRT; 186 AA.
AC 08YGD7
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
GN Hypothetical protein BME11222.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-16M / ATCC 23456 / BIOTYPE 1;
RX MEDLINE-20020109; PubMed-11756688;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mijer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Ivanovski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselhorn R., Kyrpides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL: AE009561; AAL52403.1;-
KW Hypothetical protein; Complete proteome;
SQ SEQUENCE 186 AA; 20709 MW; 82199539C204A294 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 16; Length 186;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 106 EVAKIV 112
Db 16 EVAKIV 22

RESULT 15
O9SAG6
ID 09SAG6 PRELIMINARY; PRT; 213 AA.
AC 09SAG6
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
GN F23A5.30 protein (Hypothetical 24.2 kDa protein).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV, COLUMBIA;
RA Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S., Lee J.M., Li J.,
RA Gonzalez A., Liu A., Liu K., Sakano H., Koo T., Pham P., Vaysberg M.,
RA Hong B., Chin C., Choi E., Chlou J., Altafi H., Brooks S., Chao Q.,
RA Conn L., Conway A., Hansen N., Johnson-Hopson C., Khan S., Kim C.,
RA Lam B., Nguyen M., Palm C., Shinn P., Tambunga G., Davis R.W.,
RA Ecker J.R., Federspiel N.A., Theologis A.;
RT "Arabidopsis thaliana chromosome 1 BAC F23A5 sequence.";
RN Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
[2]

RP SEQUENCE FROM N.A.
RC STRAIN-CV, COLUMBIA;
RA Theologis A.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-CV, COLUMBIA;
RA Theologis;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Etgu P., Lee J.M.,
RA Toriumi M., Yu G., Brooks S., Chao Q., Chen H., Karlin-Neumann G.,
RA Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shinn P.,
RA Southwick A., Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene F23A5.30 (GI:6503306).";
RN Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamuya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene At1g980940 (GI:15220893).";
RN Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
[6]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamuya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones.";
RN Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
[7]
RP Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC011713; AAF14682.1;-
DR EMBL: AF332462; AAG48825.1;-
DR EMBL: AY063820; AAL36176.1;-
DR EMBL: AY091385; AAM14324.1;-
KW Hypothetical protein.
SQ SEQUENCE 213 AA; 24243 MW; BF27D52F10FC274F CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 10; Length 213;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 83 DKAPEAV 89
Db 44 DKAPEAV 50

RESULT 16
O9V3Y0
ID 09V3Y0 PRELIMINARY; PRT; 215 AA.
AC 09V3Y0
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
GN BG:DS00929.2 protein (AT29831p).
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;
 MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrell J.F., Agbayani A., An H.-J., Andrews-Plamkoch C., Baldwin D.,
 RA Ballaw R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bottler P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spler E.C., Spradling A.C., Stapleton M., Stong R., Sun E.,
 RA Svitksas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Welstock G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao X., Zhang L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 [12]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE-99403001; PubMed-10471707;
 RA Ashburner M., Mistra S., Roote J., Lewis S.E., Blazer R., Davis T.,
 RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
 RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
 RA Palazolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitehead K.,
 RA Celniker S., Rubin G.M.;
 RT "An exploration of the sequence of a 2.9-Mb region of the genome of
 RT *Drosophila melanogaster*: the *Adh* region.";
 RL Genetics 153:179-219(1999).
 [13]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazer R.G.,
 RA Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
 RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
 RA Houston K.A., Hummasti S.R., Karia K., Kearney L., Kim E., Lee B.,
 RA Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
 RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequelita A.,
 RA Setchell H., Sait E., Svitksas R.R., Wan K.H., Welbush T., Zhang R.,
 RA Zieran L.T., Rubin G.M.;
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 [14]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarini H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AE003646; AAF53427.1; -;
 DR EMBL: AE003411; AAF44906.1; -;
 DR EMBL: AY070830; AAL48452.1; -;
 DR HSSP: P80144; 2MYO.
 DR FlyBase: FBgn0028535; BG:DS00929.2.
 DR InterPro: IPR002110; ANK.
 DR Pfam: PF00023; ank; 3.
 DR SMART: SM00248; ANK; 2.
 DR PROSITE: PSS0088; ANK_REPEAT; 2.
 DR PROSITE: PSS0297; ANK_REPEAT_REGION; 1.
 KW ANK repeat; Repeat.
 SQ SEQUENCE 215 AA; 24244 MW; 4638BE441733ECB7 CRC64;
 Query Match 5.2%; Score 7; DB 5; Length 215;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 100 NPNARTE 106
 Db 107 NPNARTE 113
 RESULT 17
 Q8T53 PRELIMINARY; PRT; 222 AA.
 ID 08T533
 AC 08T533;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Adenylate kinase (Fragment).
 OS Trypanosoma brucei.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5691.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nazaya E.-S., Kabiri M., Becker K., Steverding D.;
 RT "Adenylate kinase of *Trypanosoma brucei*: Molecular cloning, expression
 RT and characterization.";
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF497801; AAM18618.1; -;
 KW Kinase.
 FT NON_TER 1
 SQ SEQUENCE 222 AA; 24759 MW; 3A19ABC7EB951C6A CRC64;
 Query Match 5.2%; Score 7; DB 5; Length 222;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 105 TEVAOKI 111
 Db 73 TEVAOKI 79
 RESULT 18
 Q9O348 PRELIMINARY; PRT; 233 AA.
 ID 09O348
 AC 09O348;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Polyprotein (Fragment).
 GN POL.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MN5707;
 RX MEDLINE-20283758; PubMed-10823789;
 RA Devereux H.L., Loveday C., Youle M., Sablin C.A., Burke A.,
 RA Johnson M.A.;
 RT "Reduction in human immunodeficiency virus type 1 mutations associated
 RT with drug resistance after initiating new therapeutic regimens in

```
RT   pretreated patients."
RL   J. Infect. Dis. 181:1804-1807(2000).
DR   EMBL; AF210537; AAF18232.1; -.
DR   InterPro; IPR000477; RVTSE.
DR   Pfam; PF00078; rvt; 1.
KW   RNA-directed DNA polymerase.
FT   NON_TER 1
FT   NON_TER 233
SQ   SEQUENCE 233 AA; 27150 MW; 6FE4AB1B6C17FD56 CRC64;

Query Match
Best Local Similarity 5.2%; Score 7; DB 15; Length 233;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TATFLAS 19
    |||||
DB 121 TATFLAS 127

RESULT 19
ID 090877 PRELIMINARY; PRT; 246 AA.
AC 090877;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE EV-21 envelope (fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=92141069; Pubmed=1685775;
RA Levin I., Smith E.J.;
RT "Sequence analysis of avian endogenous virus ev21 integration site and
RT identification of an associated sex linked CRI element."
RL Poul. Sci. 70:1948-1956(1991).
DR EMBL; X54094; CAA38028.1; -.
DR InterPro; IPR005166; Avian_gp85.
DR Pfam; PF03708; Avian_gp85; 1.
FT NON_TER 1
SQ SEQUENCE 246 AA; 27040 MW; C823BA4254F2AC0D CRC64;

Query Match
Best Local Similarity 5.2%; Score 7; DB 13; Length 246;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 ANITTSLS 39
    |||||
DB 92 ANITTSLS 98

RESULT 20
ID 085408 PRELIMINARY; PRT; 246 AA.
AC 085408;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Avian endogenous virus RAV-0 env gene 3'-region for C-terminal region
DE of envelope glycoproteins gp85 and gp37 (Fragment).
OS Avian endogenous rous-associated virus-0 (EV-2) (Avian retrovirus
OS RAV-0).
OC Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.
OX NCBI_TaxID=11949;
RN [1]
RP SEQUENCE FROM N.A.
RA Tikhonenko A.T.;
RL Submitted (May-1988) to the EMBL/GenBank/DBJ databases.
RN [2]
OX SEQUENCE FROM N.A.
```

```
RX MEDLINE=90266494; Pubmed=2161159;
RA Tikhonenko A.T., Lomovskaya O.L.;
RT "Avian endogenous provirus (ev-3) env gene sequencing: implications
RT for pathogenic retrovirus origination."
RL Virus Genes 3:251-258(1990).
DR EMBL; X07818; CAA30677.1; -.
DR InterPro; IPR005166; Avian_gp85.
DR Pfam; PF03708; Avian_gp85; 1.
FT NON_TER 1
SQ SEQUENCE 246 AA; 26871 MW; 4AAFAE21E70CC05B CRC64;

Query Match
Best Local Similarity 5.2%; Score 7; DB 15; Length 246;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 ANITTSLS 39
    |||||
DB 92 ANITTSLS 98

RESULT 21
ID 09XDX8 PRELIMINARY; PRT; 255 AA.
AC 09XDX8;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE A-type carbonic anhydrase (EC 4.2.1.1).
GN ACAP.
OS Rhodopsudomonas palustris.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Rhodopsudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RA SPRAIN-NO. 7.
RX MEDLINE=20519261; Pubmed=11065374;
RA Puskas L.G., Inui M., Zahn K., Yukawa H.;
RT "A periplasmic, alpha-type carbonic anhydrase from Rhodopsudomonas
RT palustris is essential for bicarbonate uptake."
RL Microbiology 146:2957-2966(2000).
DR EMBL; AB022175; BAA82053.1; -.
DR HSSP; Q50940; IKOQ.
DR InterPro; IPR001148; Euk_COanh.
DR Pfam; PF00194; carb_anhydrase; 1.
DR ProDom; PD000865; Euk_COanh.2.
DR PROSITE; PS00162; Euk_CO2_ANHYDRASE; UNKNOWN_1.
KW lyase.
SQ SEQUENCE 255 AA; 27303 MW; 26261B8E4F56975 CRC64;

Query Match
Best Local Similarity 5.2%; Score 7; DB 2; Length 255;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 AAINPNA 103
    |||||
DB 183 AAINPNA 189

RESULT 22
ID 016217 PRELIMINARY; PRT; 277 AA.
AC 016217;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 32.0 kDa protein.
GN F17A9.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita; Rhabdilitidae;
OC Rhabdilitae; Peloderitinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
```

RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE-99069613; Pubmed-9851916;
 RA None;
 RT "Genome sequence of the nematode *C. elegans*: a platform for
 RT investigating biology. The *C. elegans* Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Wohlmann P., Sammons L., Rohlfing T., Gillingham B.;
 RT "The sequence of *C. elegans* cosmid F17A9.";
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF016417; AAB65278.1;
 DR InterPro; IPR002900; DUF38.
 DR Pfam; PF01827; DUF38; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 277 AA; 31967 MW; 77AB19751AB03B80 CRC64;

Query Match 5.2%; Score 7; DB 5; Length 277;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 90 SLTSLT 96
 |||||
 DB 184 SLTSLT 190

RESULT 23

ID O9K059 PRELIMINARY; PRT; 302 AA.
 AC O9K059;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Tyra protein.
 GN VC2145.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE-20406833; Pubmed-10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Esmolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nieman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
 RT cholerae.";
 RL Nature 406:477-483(2000).
 DR EMBL; AE004287; AAF95290.1;
 DR TIGR; VC2145;
 KM Complete proteome.
 SQ SEQUENCE 302 AA; 33496 MW; B45BAC93888F2F7E CRC64;

Query Match 5.2%; Score 7; DB 16; Length 302;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 110 KIVRHSL 116
 |||||
 DB 227 KIVRHSL 233

RESULT 24
 ID O93892 PRELIMINARY; PRT; 320 AA.
 AC O93892;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE M02B1.3 protein.
 GN M02B1.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RL Lightning J.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99069613; Pubmed-9851916;

Query Match 5.2%; Score 7; DB 5; Length 320;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 NLTSLI 40
 |||||
 DB 91 NLTSLI 97

RESULT 25

ID O9RCH0 PRELIMINARY; PRT; 362 AA.
 AC O9RCH0;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE Tyra.
 GN VC2145.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-569B;
 RA Schroeder I.C., Coyne V.E.;
 RT "Cloning and characterization of the melanin biosynthetic genes from
 RT *Vibrio cholerae* 569B.";
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U62056; AAF21219.1;
 SQ SEQUENCE 362 AA; 39980 MW; 7A3A18F8A2D0BE7B CRC64;

Query Match 5.2%; Score 7; DB 2; Length 362;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 110 KIVRHSL 116
 |||||
 DB 227 KIVRHSL 233

RESULT 26
 ID O96X43 PRELIMINARY; PRT; 389 AA.
 AC O96X43;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Ure2p.
 GN URE2.
 OS Kluyveromyces lactis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
 OX NCBI_TaxID=28985;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fernandez-Bellot E., Baudin-Ballieu A., Cullin C.;
 RT "Pilon characteristics of the URE2 protein of various yeast species."
 RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF260776; AAK51642.1; -;
 DR InterPro: IPR004046; GST_Cterm.
 DR InterPro: IPR004045; GST_Nterm.
 DR Pfam: PF00043; GST_C; 1.
 DR Pfam: PF02798; GST_N; 1.
 SO SEQUENCE 389 AA; 44395 MW; CF4C33F844FBFB65 CRC64;

Query Match 5.2%; Score 7; DB 3; Length 389;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 AINPNAR 104
 DB 193 AINPNAR 199

RESULT 27
 O9GSG9 PRELIMINARY; PRT; 399 AA.
 AC O9GSG9;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
 DE 01-MAR-2002 (TREMBLrel. 20, last annotation update)
 GN Wiscott-aldrich syndrome protein.
 OS Dictyostellium discoideum (slime mold).
 OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chung C.Y., Han Y.-H., Firtel R.A.;
 RT "Regulation of actin cytoskeleton by WASP during Dictyostellium
 chemotaxis."
 RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF305060; AAG24442.1; -;
 DR InterPro: IPR000095; Pakbox/Rhobndng.
 DR InterPro: IPR002965; P_trich_extensn.
 DR InterPro: IPR000697; Ranbp1_MASP.
 DR InterPro: IPR001960; WH1.
 DR Pfam: PF00786; PBD; 1.
 DR Pfam: PF00568; WH1; 1.
 DR Pfam: PF02205; WH2; 1.
 DR PRINTS: PRO1217; PRICHEXTENS.
 DR SMART: SM00285; PBD; 1.
 DR SMART: SM00461; WH1; 1.
 DR SMART: SM00246; WH2; 1.
 DR PROSITE: PSS0108; GHD; 1.
 SO SEQUENCE 399 AA; 42446 MW; AOB8BFCAR365E317 CRC64;

Query Match 5.2%; Score 7; DB 5; Length 399;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 96 TAAINPN 102
 DB 129 TAAINPN 135

RESULT 28

O8U7N5
 ID O8U7N5 PRELIMINARY; PRT; 436 AA.
 AC O8U7N5;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, last sequence update)
 DE ABC transporter, membrane spanning protein.
 GN RSPE OR AT04414 OR AGR_L_906.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-21608550; PubMed-11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D., St.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet M., Grant C.,
 RA Kuyavlin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan M., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 C58."
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-21608551; PubMed-11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard C.,
 RA Ourullo B., Goldman B.S., Cao Y., Askenazi M., Helling M., Mullin L.,
 RA Houmlel K., Gordon J., Vaudin M., Tarrhouk O., Bpp A., Liu F.,
 RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Markelz B.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58."
 RL Science 294:2323-2328(2001).
 DR EMBL; AE009370; AAL45208.1; ALT_INTT.
 DR EMBL; AE008244; AAK89025.1; -;
 KW Complete proteome.
 SO SEQUENCE 436 AA; 48511 MW; 9C47850D301E9863 CRC64;

Query Match 5.2%; Score 7; DB 16; Length 436;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 71 AKISETA 77
 DB 238 AKISETA 244

RESULT 29
 O9XL11 PRELIMINARY; PRT; 459 AA.
 ID O9XL11;
 AC O9XL11;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)
 DE 01-MAR-2002 (TREMBLrel. 20, last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
 OS Falco peregrinus.
 OC Mitochondrion.
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Falconiformes; Falconidae; Falco.
 OX NCBI_TaxID=8954;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mindell D.P., Sorenson M.S., Huddleston C.J., Miranda H.C. Jr.,
 RA Knight A., Sawchuk S.J., Turi T.;
 RT "Phylogenetic relationships among and within select avian orders based
 on mitochondrial DNA."

RL (In) Mindell D.F. (eds.);
 RL AVIAN MOLECULAR EVOLUTION AND SYSTEMATICS, pp.211-245,
 RL Academic Press (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-98393707; PubMed-9724766;
 RA Mindell D.P., Sorenson M.D., Dimcheff D.E.;
 RT "Multiple independent origins of mitochondrial gene order in birds."
 RL Proc. Natl. Acad. Sci. U.S.A. 95:10693-10697(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Mindell D.P., Sorenson M.D., Dimcheff D.E., Hasegawa M., Ast J.C.,
 RA Yuri T.;
 RT "Interordinal relationships of birds and other reptiles based on whole
 RT mitochondrial genomes."
 RL Syst. Biol. 48:138-152(1999).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Sorenson M.D., Dimcheff D.E., Ast J.C., Yuri T., Mindell D.P.;
 RT "Complete mitochondrial DNA sequences for five birds and a turtle."
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
 DR EMBL: AF090338; AAD32501.1; -;
 DR InterPro: IPR001750; Oxidored_q1.
 DR InterPro: IPR002260; Oxidored_q1.
 DR Pfam: PF00361; oxidored_q1; 1.
 DR Pfam: PF01059; oxidored_q5_N; 1.
 KM Mitochondrion; NAD: Oxidoreductase; Ubiquinone.
 SQ SEQUENCE 459 AA; 51109 MW; FCD1BPIBC01AD0 CRC64;

Query Match 5.2%; Score 7; DB 8; Length 459;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 TATLMT 14
 DB 398 TATLMT 404

RESULT 30
 026204 PRELIMINARY; PRT; 499 AA.
 AC 026204;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein MTH101.
 GN MTH101.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DELTA H;
 RX MEDLINE-98037514; PubMed-9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Polther B., Qiu D.,
 RA Spadefora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Sater H., Patwell D., Prabhakar S.,
 RA McDougall S., Shiner G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT deltaH: functional analysis and comparative genomics."
 RL J. Bacteriol. 179:7135-7155(1997).
 DR EMBL: AE000800; AAB84600.1; -;
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 499 AA; 56132 MW; B0DD36CFC3D9A4 CRC64;

Query Match 5.2%; Score 7; DB 17; Length 499;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 VTATLMT 13
 DB 454 VTATLMT 460

RESULT 31
 08T0Y6 PRELIMINARY; PRT; 501 AA.
 ID 08T0Y6;
 AC 08T0Y6;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein MA1400.
 GN MA1400.
 OS Methanosarcina acetivorans.
 OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE-21929760; PubMed-11932238;
 RA Galagan J.E., Nusbäum C., Roy A., Endrizzi M.G., Macdonald P.,
 RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Alnoor D., Brown A.,
 RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
 RA Linton L., McEwan P., McKernan K., Talamas J., Filrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame P.A., Guss A.M.,
 RA Haderich R., Ingram-Smith C., Kueltner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.;
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 RT and physiological diversity."
 RL Genome Res. 12:532-542(2002).
 DR EMBL: AE010809; AAM04816.1; -;
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 501 AA; 56015 MW; 35FAA2ED9A8E0908 CRC64;

Query Match 5.2%; Score 7; DB 17; Length 501;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 VTATLMT 13
 DB 457 VTATLMT 463

RESULT 32
 007453 PRELIMINARY; PRT; 577 AA.
 ID 007453;
 AC 007453;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE ENV polyprotein [CONTAINS: coat protein GP85; coat protein GP37].
 GN ENV.
 OS Rous sarcoma virus.
 OC Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.
 OX NCBI_TaxID=11886;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PR-RSV-C;
 RX MEDLINE-93254440; PubMed-8387633;
 RA Kashba V.I., Kavan V.M., Ryndich A.V., Lazurkevich Z.V., Zubak S.V.,
 RA Popov S.V., Dostalova V., Hlozaneck I.;
 RT "Complete nucleotide sequence of Rous sarcoma virus variants adapted
 RT to duck cells."
 RL Mol. Biol. (Mosk) 27:436-450(1993).
 CC -1- MISCELLANEOUS: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE
 CC PROTEINS.

DR EMBL: X68524; CAA48536.1; -
 DR InterPro: IPR005166; Avian_gp85.
 DR Pfam: PF03708; Avian_gp85; 1.
 KW Coat protein; Polyprotein.
 FT CHAIN 39 577 COAT PROTEIN GP85.
 FT CHAIN 380 577 COAT PROTEIN GP37.
 SQ SEQUENCE 577 AA; 62591 MW; F014A4EFB8254B44 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 7; DB 15; Length 577;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39
 |||||
 DB 430 ANLTSL 436

RESULT 33
 003819 PRELIMINARY; PRT; 595 AA.

ID 003819
 AC 003819; PRELIMINARY; PRT; 595 AA.
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE ENV polyprotein [CONTAINS: coat protein GP85; coat protein GP37]
 DE (Fragment).
 GN ENV.
 OS Rous sarcoma virus.
 OC Viruses; Retrov. viruses; Retroviridae; Alpharetrovirus.

OC NCBI_Taxid=11886;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zubak S.V.;
 RL Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 398-580 FROM N.A.
 RX MEDLINE=90114140; Pubmed=2558292;
 RA Ryndlich A.V., Kashuba V.I., Kavan V.M., Zubak S.V., Ghozhanek I.;
 RT "The family of env genes of avian retroviruses: molecular analysis of
 Rous sarcoma virus adapted to duck cells.";
 RL Mol. Biol. (Mosk) 23:1355-1363(1989).
 CC -1 MISCELLANEOUS: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE
 PROTEINS.

DR EMBL: X51862; CAA36155.1; -
 DR InterPro: IPR005166; Avian_gp85.
 DR Pfam: PF03708; Avian_gp85; 1.
 KW Coat protein; Polyprotein.
 FT CHAIN 1 1 COAT PROTEIN GP85.
 FT CHAIN 398 595 COAT PROTEIN GP37.
 FT CHAIN 57 595 COAT PROTEIN GP85.
 SQ SEQUENCE 595 AA; 64540 MW; 139B3C572B18E364 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 7; DB 15; Length 595;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39
 |||||
 DB 448 ANLTSL 454

RESULT 34
 083129 PRELIMINARY; PRT; 600 AA.

ID 083129
 AC 083129; PRELIMINARY; PRT; 600 AA.
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE ENV protein.
 GN ENV.
 OS Avian myeloblastosis-associated virus 1/2.
 OC Viruses; Retrov. viruses; Retroviridae; Avian type C retroviruses.

OC NCBI_Taxid=31668;
 RX SEQUENCE FROM N.A.

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93331743; Pubmed=8393249;
 RA Joliet V., Borouh K., Lasserre F., Crochet J., Dambine G.,
 RA Smith R.E., Perbal B.;
 RT "Pathogenic potential of myeloblastosis-associated virus: implication
 of ENV proteins for osteopetrosis induction.";
 RL Virology 195:812-819(1993).
 DR EMBL: L10923; AAA46300.1; -
 DR InterPro: IPR005166; Avian_gp85.
 DR Pfam: PF03708; Avian_gp85; 1.
 SQ SEQUENCE 600 AA; 65393 MW; F6764AD2F4D7D1C3 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 7; DB 15; Length 600;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39
 |||||
 DB 453 ANLTSL 459

RESULT 35
 090VD8 PRELIMINARY; PRT; 601 AA.

ID 090VD8
 AC 090VD8; PRELIMINARY; PRT; 601 AA.
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Recombinant envelope polypeptide (Fragment).
 GN ENV.
 OS Avian leukosis virus.
 OC Viruses; Retrov. viruses; Retroviridae; Alpharetrovirus.

OC NCBI_Taxid=11864;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20478055; Pubmed=11021992;
 RA Luplant B., Hunt H., Silva R., Fadly A.;
 RT "Identification and characterization of recombinant subgroup J avian
 RT leukosis viruses (ALV) expressing subgroup A ALV envelope.";
 RL Virology 276:37-43(2000).
 DR EMBL: AF257655; AK57473.1; -
 DR InterPro: IPR005166; Avian_gp85.
 DR Pfam: PF03708; Avian_gp85; 1.
 KW Envelope protein; Polyprotein.
 FT CHAIN 1 1
 SQ SEQUENCE 601 AA; 65655 MW; 6C7BB826B5B94B6 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 7; DB 15; Length 601;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39
 |||||
 DB 447 ANLTSL 453

RESULT 36
 090VD7 PRELIMINARY; PRT; 601 AA.

ID 090VD7
 AC 090VD7; PRELIMINARY; PRT; 601 AA.
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Recombinant envelope polypeptide (Fragment).
 GN ENV.
 OS Avian leukosis virus.

OC Viruses; Retrov. viruses; Retroviridae; Alpharetrovirus.
 OC NCBI_Taxid=11864;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ADOL 6803A;

RA MEDLINE-20478055; PubMed-11021992;
 RA Lupiani B., Hunt H., Silva R., Faddy A.;
 RT "Identification and characterization of recombinant subgroup J avian
 RT Leukosis viruses (ALV) expressing subgroup A ALV envelope.";
 RL Virology 276:37-43(2000).
 DR EMBL: AF257656; AAK57474.1; -;
 DR InterPro: IPR005166; Avian_gp85.
 DR Pfam: PF03708; Avian_gp85; 1.
 KW Envelope protein; Polyprotein.
 FT NON_TER 1
 SQ SEQUENCE 601 AA; 65645 MW; 12E9D417B5FB6887 CRC64;

Query Match 5.2%; Score 7; DB 15; Length 601;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39
 DB 447 ANLTSL 453

RESULT 37
 ID 090VD6 PRELIMINARY; PRT; 601 AA.
 AC 090VD6;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Recombinant envelope polyprotein (Fragment).
 GN ENV.
 OS Avian leukosis virus.
 OC Viruses; Retrovirdae; Retroviridae; Alpharetrovirus.
 OX NCBI_TaxID-11864;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RAV 0-A1;
 RX MEDLINE-20478055; PubMed-11021992;
 RA Lupiani B., Hunt H., Silva R., Faddy A.;
 RT "Identification and characterization of recombinant subgroup J avian
 RT Leukosis viruses (ALV) expressing subgroup A ALV envelope.";
 RL Virology 276:37-43(2000).
 DR EMBL: AF257657; AAK57475.1; -;
 DR InterPro: IPR005166; Avian_gp85.
 DR Pfam: PF03708; Avian_gp85; 1.
 KW Envelope protein; Polyprotein.
 FT NON_TER 1
 SQ SEQUENCE 601 AA; 65645 MW; 12E9D417B5FB6887 CRC64;

Query Match 5.2%; Score 7; DB 15; Length 601;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39
 DB 447 ANLTSL 453

RESULT 38
 ID 090VD5 PRELIMINARY; PRT; 601 AA.
 AC 090VD5;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Recombinant envelope polyprotein (Fragment).
 GN ENV.
 OS Avian leukosis virus.
 OC Viruses; Retrovirdae; Retroviridae; Alpharetrovirus.
 OX NCBI_TaxID-11864;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ADOL 5701A DELTA;
 RX MEDLINE-20478055; PubMed-11021992;

RA Lupiani B., Hunt H., Silva R., Faddy A.;
 RT "Identification and characterization of recombinant subgroup J avian
 RT Leukosis viruses (ALV) expressing subgroup A ALV envelope.";
 RL Virology 276:37-43(2000).
 DR EMBL: AF257658; AAK57476.1; -;
 DR InterPro: IPR005166; Avian_gp85.
 DR Pfam: PF03708; Avian_gp85; 1.
 KW Envelope protein; Polyprotein.
 FT NON_TER 1
 SQ SEQUENCE 601 AA; 65655 MW; 6C7BB826B5B94B6 CRC64;

Query Match 5.2%; Score 7; DB 15; Length 601;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39
 DB 447 ANLTSL 453

RESULT 39
 ID 064984 PRELIMINARY; PRT; 601 AA.
 AC 064984;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE ENV-PR95 polyprotein precursor.
 GN ENV.
 OS Rous sarcoma virus (strain Prague C).
 OC Viruses; Retrovirdae; Retroviridae; Alpharetrovirus.
 OX NCBI_TaxID-11888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PRAGUE;
 RX MEDLINE-83155662; PubMed-6299578;
 RA Schwartz D.E., Tizard R., Gilbert W.;
 RT "Nucleotide sequence of Rous sarcoma virus.";
 RL Cell 32:853-869(1983).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PRAGUE;
 RX MEDLINE-85124605; PubMed-2982497;
 RA Broome S., Gilbert W.;
 RT "Rous sarcoma virus encodes a transcriptional activator.";
 RL Cell 40:537-546(1985).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Petropoulos C.J.;
 RL (in) Coffin J.M. (eds.);
 RL Retroviruses, pp.757-757, Cold Spring Harbor Laboratory Press,
 RL New York (1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Chappey C.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: J02342; AAB59934.1; -;
 DR EMBL: AF033808; AAC82562.1; -;
 DR InterPro: IPR005166; Avian_gp85.
 DR Pfam: PF03708; Avian_gp85; 1.
 KW Polyprotein.
 FT CHAIN 1 175 POTENTIAL.
 FT CHAIN 63 601 POTENTIAL.
 FT CHAIN 404 601 POTENTIAL.
 SQ SEQUENCE 601 AA; 65316 MW; A3FD5DABD1A832C CRC64;

Query Match 5.2%; Score 7; DB 15; Length 601;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39
 DB 454 ANLTSL 460

RESULT 40

083132
 ID 083132 PRELIMINARY; PRT; 604 AA.
 AC 083132;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Env Protein.
 GN ENV.
 OS Avian myeloblastosis-associated virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Avian type C retroviruses.
 OX NCBI_TaxID=31669;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93331743; PubMed=8393249;
 RA Joliet V., Boroughs K., Lasserre F., Crochet J., Dambrine G.,
 RA Smith R.E., Perbal B.;
 RT "Pathogenic potential of Myeloblastosis-Associated Virus: Implication
 of ENV Proteins for Osteopetrosis Induction."
 RL Virology 195:812-819(1993).
 DR EMBL; L10922; AAA46303.1; -;
 DR InterPro; IPR005166; Avian_gp85.
 DR Pfam; PF03708; Avian_gp85; 1.
 SO SEQUENCE 604 AA; 65930. MW; 8536766E47F3FC3 CRC64;

Query Match

Best Local Similarity 5.2%; Score 7; DB 15; Length 604;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 ANLTSL 39
 |||||
 DB 457 ANLTSL 463

Search completed: April 28, 2003, 16:15:00
 Job time : 130 secs